

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 08:17:12 ; Search time 76 Seconds
(without alignments)
3810.690 Million cell updates/sec

Title: US-09-662-454-3
Perfect score: 16009
Sequence: 1 MNTNPKPKRKTNTNRPPQ.....FPLCLLLSVGVGVIYLLPMR 3010

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15367	96.0	3010	1	GNMWVCJ genome polyprotein
2	15253	95.3	3010	1	GNMWVCJ genome polyprotein
3	15236	95.2	3010	1	A45573 genome polyprotein
4	15062	94.1	3010	1	GNMWVCJ genome polyprotein
5	15052	94.0	3010	1	GNMWVCJ genome polyprotein
6	14119.5	88.2	3011	1	GNMWVCJ genome polyprotein
7	14008.5	87.5	3011	1	GNMWVCJ genome polyprotein
8	13935.5	87.0	3011	1	GNMWVCJ genome polyprotein
9	12496	78.1	3014	1	JC5620 genome polyprotein
10	12036.5	75.2	3033	1	JQ1303 genome polyprotein
11	11972.5	74.8	3033	1	GNMWVCJ genome polyprotein
12	3977	24.8	782	2	S18031 genome polyprotein
13	3952	24.7	782	2	S18032 genome polyprotein
14	3942	24.6	782	2	S18032 genome polyprotein
15	3922	24.5	782	2	S18032 genome polyprotein
16	3916	24.5	787	2	PN0677 hypotheical prote
17	3473	21.7	876	2	PC2219 polypeptide - hepa
18	3051	19.1	640	2	JQ1584 genome polyprotein
19	3023	18.9	716	2	JQ1366 genome polyprotein - hepa
20	2674	16.7	513	2	PC1284 genome polyprotein
21	2581.5	16.1	523	2	JQ1926 polyprotein - hepa
22	2577.5	16.1	520	2	JQ1925 polyprotein - hepa
23	2513	15.7	550	2	JH0711 genome polyprotein
24	2426	15.2	874	2	JQ0883 genome polyprotein
25	2417	15.1	874	2	JQ0881 genome polyprotein
26	2393	14.9	513	2	A44150 structural protein
27	2387	14.9	492	2	PS0326 polyprotein - hepa
28	2286.5	14.3	3005	2	T08841 polyprotein - dour
29	2251	14.1	441	2	S12707 genome polyprotein

30	2131	13.3	2970	2	T08839 polyprotein - marm
31	2103.5	13.1	492	2	S41288 genome polyprotein
32	2039	12.7	415	2	PC4407 envelope protein -
33	1907.5	11.9	386	2	ATPase/RNA helicas
34	1892	11.8	365	2	JQ0879 NS5 protein - hepa
35	1880.5	11.7	369	2	S21471 genome polyprotein
36	1779	11.1	411	2	PC2061 genome polyprotein
37	1751	10.9	411	2	JQ0880 genome polyprotein
38	1741	10.9	365	2	JQ0880 NS5 protein - hepa
39	1576	9.8	322	2	JN0265 genome polyprotein
40	1531	9.6	315	2	PN0011 envelope glycoprot
41	1517	9.5	315	2	PS0164 envelope glycoprot
42	1491	9.3	315	2	PS0165 genome polyprotein
43	1461	9.1	350	2	PS5631 genome polyprotein
44	1161	7.3	234	2	S32742 genome polyprotein
45	1135.5	7.1	235	2	S32747 genome polyprotein

ALIGNMENTS

RESULT 1

GNMWVCJ genome polyprotein - hepatitis C virus (strain J)

N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C/Accession: A39253; PS0086

R:Kato, N.; Hijikata, M.; Ohtsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimoto

Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990

A>Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients

A/Reference number: A39253; PMID:9108550; PMID:2175903

A/Accession: A39253

A/Molecule type: genomic RNA

A/Residues: 1-3010 <KAT>

A/Cross-references: UNIPROT:P26662; UNIPARC:UPI0000131E28; GB:ID90208; NID:9221610; PIDN:1

R:Kato, N.; Ohkoshi, S.; Shimotohno, K.

Proc. Jpn. Acad. 65B, 219-223, 1989

A>Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari

A/Reference number: PS0085

A/Accession: PS0086

A/Molecule type: genomic RNA

A/Residues: 2650-2707 <KAT>

A/Cross-references: UNIPARC:UPI00000F7365

A/Experimental source: Japanese isolate

C/Comment: The cleavage sites of this polyprotein have not been determined.

C/Superfamily: hepatitis C virus genome polyprotein

C/Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin

F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <EPM>

F:199-389/Product: major envelope protein E #status predicted <MEE>

F:730-1006/Product: nonstructural protein NS1 #status predicted <NS1>

F:1230-1237/Product: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1316-1319/Region: DEHX motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F:196,209,224,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2

Query Match 96.0%; Score 15367; DB 1; Length 3010; Best Local Similarity 95.2%; Pred. No. 0; Mismatches 64; Indels 0; Gaps 0;

QY 1 MSTNPKRKTNTNRPPQVKKFPGGQIVGVYLLPRRGRPRGVRATRRASERSQPRG 60
DB 1 MSTNPKRKTNTNRPPQVKKFPGGQIVGVYLLPRRGRPRGVRATRRATRSERSQPRG 60

QY 61 RRQPIKARPEGRMAQPGYVWPLVYNGSLGMAWLLSPGSRPSSWQPTDPRRSRLG 120
DB 61 RRQPIKARPEGRMAQPGYVWPLVYNGSLGMAWLLSPGSRPSSWQPTDPRRSRLG 120

Db 61 RROPIPKARRPBRGRTAQQPCYWPPLYGNBMCMAGMLLSBRCGRBWCPTDPRRRNRNG 120
QY 121 KYIDTLTCGFADIMGYIPLVGAPLGGABALAHGVLEDEGVNATGNLPGCSFSIFLLA 180
Db 122 KYIDTLTCGFADIMGYIPLVGAPLGGABALAHGVLEDEGVNATGNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASAEVRNVSGIYHVTNDCSNSIYEEADVIYMTPGCVPCVQBSNRRCAV 240
Db 181 LLSCLTTPASAEVRNVSGIYHVTNDCSNSIYEEADVIYMTPGCVPCVQBSNRRCAV 240
QY 241 ALTPTLAANASVPTTIRRHVDLIVGTAFCGAMVGLCSIFLVSLTFSPRRHET 300
Db 241 ALTPTLAANASVPTTIRRHVDLIVGTAFCGAMVGLCSIFLVSLTFSPRRHET 300
QY 301 VODNCNSIYGVHSGHRMAMDMMMNSPTTALVVSOLLRIPOAVDMVAGAHGVLAGLA 360
Db 301 VODNCNSIYGVHSGHRMAMDMMMNSPTTALVVSOLLRIPOAVDMVAGAHGVLAGLA 360
QY 361 YYSWGNMAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTLSFSSGASOKIQLVNTNGSW 420
Db 361 YYSWGNMAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTLSFSSGASOKIQLVNTNGSW 420
QY 421 HINRTALNNDLSQTCGFPAALFYAHKFNSSGCEBMASSCRPIDMFAQSGCPIYTTKPNSS 480
Db 421 HINRTALNNDLSQTCGFPAALFYAHKFNSSGCEBMASSCRPIDMFAQSGCPIYTTKPNSS 480
QY 481 DORPYCMHAPPCGVVPASOVCGPYCFTPSPVVVGTTDRSGVPTYSWGENETDMLN 540
Db 481 DORPYCMHAPPCGVVPASOVCGPYCFTPSPVVVGTTDRSGVPTYSWGENETDMLN 540
QY 541 NTRPPOGNWPGCTWNSSTGFTKTCGGRPCNIGGVNKRILI CPTDCRRKHPBEATYTKCSG 600
Db 541 NTRPPOGNWPGCTWNSSTGFTKTCGGRPCNIGGVNKRILI CPTDCRRKHPBEATYTKCSG 600
QY 601 PWLTPECLVDYPRYLHYPCTLNFSI EKYRMVYGVGEHRLNAA CMTREGRCLBRRDS 660
Db 601 PWLTPECLVDYPRYLHYPCTLNFSI EKYRMVYGVGEHRLNAA CMTREGRCLBRRDS 660
QY 661 ELSPLLSTTEMQILPCAFETTLPALSTGLIHLHONIVDOYLYGVGSAFSPAIKMEYIL 720
Db 661 ELSPLLSTTEMQILPCAFETTLPALSTGLIHLHONIVDOYLYGVGSAFSPAIKMEYIL 720
QY 721 LFLFLIADARVCA CMMMLIIAQAEBALENLVVLNAAVAGAGILSFLVFCAMYYIKG 780
Db 721 LFLFLIADARVCA CMMMLIIAQAEBALENLVVLNAAVAGAGILSFLVFCAMYYIKG 780
QY 781 RLAPGAAYAFYGVWPLLLILALPPRAYALDRMASSCGAVLVGLVFTLSYYVYFILT 840
Db 781 RLAPGAAYAFYGVWPLLLILALPPRAYALDRMASSCGAVLVGLVFTLSYYVYFILT 840
QY 841 RLIMWLQYFTRAEAHMOYWPPLANRGGRDAIILITCAVHPELIFDITKLLALILGPLM 900
Db 841 RLIMWLQYFTRAEAHMOYWPPLANRGGRDAIILITCAVHPELIFDITKLLALILGPLM 900
QY 901 VLOAGITRVPYFVRAOGLIRACMLVRKVAAGHYVQWFMKLGALITGYVYNHLTPLRDMA 960
Db 901 VLOAGITRVPYFVRAOGLIRACMLVRKVAAGHYVQWFMKLGALITGYVYNHLTPLRDMA 960
QY 961 HAGLRDLA VAEVPUVSAMETKUYTGADTACGDIIILGLPVASARGKEIFLGRADSLRG 1020
Db 961 HAGLRDLA VAEVPUVSAMETKUYTGADTACGDIIILGLPVASARGKEIFLGRADSLRG 1020
QY 1021 OGMRLIAPITAYSOQRGVLGCIITSLTGRDKNOVGEVQVNSTATOSPLATCINDCMT 1080
Db 1021 OGMRLIAPITAYSOQRGVLGCIITSLTGRDKNOVGEVQVNSTATOSPLATCINDCMT 1080
QY 1081 VYHGAOSKTIAGKGPITQYMTNVLDIVGMQAPGARSMTPCSCSSDILYVTRHADYI 1140
Db 1081 VYHGAOSKTIAGKGPITQYMTNVLDIVGMQAPGARSMTPCSCSSDILYVTRHADYI 1140
QY 1141 PYRRRDSGSLSPRPVSYLKSSGSGPLICPSGHVGVFRAAVCTRGAKAVDPIPVES 1200
Db 1141 PYRRRDSGSLSPRPVSYLKSSGSGPLICPSGHVGVFRAAVCTRGAKAVDPIPVES 1200

QY 1201 METTMSPVFTDNTSPPAVPOTFQVAHLHAPTGSGSKTKVPAAYAAQGYKVLVLPNSVA 1260
Db 1201 METTMSPVFTDNTSPPAVPOTFQVAHLHAPTGSGSKTKVPAAYAAQGYKVLVLPNSVA 1260
QY 1261 TLGFGAYSKAGIDPNIRITGRITTTGSGITYSYTKGFILADGCGSGGAYDIIIDECBS 1320
Db 1261 TLGFGAYSKAGIDPNIRITGRITTTGSGITYSYTKGFILADGCGSGGAYDIIIDECBS 1320
QY 1321 TDSITLIGITVLDQAEETAGARLVILATATPPGCVTVPHNIEIGLSNNGEIPFYGKAI 1380
Db 1321 TDSITLIGITVLDQAEETAGARLVILATATPPGCVTVPHNIEIGLSNNGEIPFYGKAI 1380
QY 1381 PIEAIKGRHLLFCHSKKKCDELAAKLTGLGINAAVYRGLDVSYI PPIGDVVVAATDAL 1440
Db 1381 PIEAIKGRHLLFCHSKKKCDELAAKLTGLGINAAVYRGLDVSYI PPIGDVVVAATDAL 1440
QY 1441 MTGFTGDPDSVIDCMTCTYQYVDESUDPTTETTTVPODAVSROGRGRTGRGSGIYR 1500
Db 1441 MTGFTGDPDSVIDCMTCTYQYVDESUDPTTETTTVPODAVSROGRGRTGRGSGIYR 1500
QY 1501 FVTPGERBEGMFDSSVLCBCEYDAGCAWYELTPAETSVRRAVLTNPGLPVCDHLEPMES 1560
Db 1501 FVTPGERBEGMFDSSVLCBCEYDAGCAWYELTPAETSVRRAVLTNPGLPVCDHLEPMES 1560
QY 1561 VFTGLTHIDAHFLSQTKQAGDNFPYLVAVYQATVCAQAQAPPSWDMQMKCLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSQTKQAGDNFPYLVAVYQATVCAQAQAPPSWDMQMKCLIRLKPTLHG 1620
QY 1621 PTPPLYLRLCAVONEVILTHPTIKYIMACMSADLEVVYTSWVLVGVYLAALAAICLITTSV 1680
Db 1621 PTPPLYLRLCAVONEVILTHPTIKYIMACMSADLEVVYTSWVLVGVYLAALAAICLITTSV 1680
QY 1681 VIVGRILISGKPAVVPDREVLVQEBDEMECASOLPYIEQOQLAEQFOKXALGLLQRTAT 1740
Db 1681 VIVGRILISGKPAVVPDREVLVQEBDEMECASOLPYIEQOQLAEQFOKXALGLLQRTAT 1740
QY 1741 KOAEAAAPVESKMPALETFMAKHMNFIISGIOYLAGISTLPGNPAIASIAFTASITSP 1800
Db 1741 KOAEAAAPVESKMPALETFMAKHMNFIISGIOYLAGISTLPGNPAIASIAFTASITSP 1800
QY 1801 LTTONTLLFNILIGSVAAQIAPPSAASAFVAGIAGAAVGSIGLKVTVLDIIAGYAGVYA 1860
Db 1801 LTTONTLLFNILIGSVAAQIAPPSAASAFVAGIAGAAVGSIGLKVTVLDIIAGYAGVYA 1860
QY 1861 GALVAFKWMSGEVPEDEULVNLLPAILSPGALVGVVCAIIRRHVGPBEGA VOMNRLI 1920
Db 1861 GALVAFKWMSGEVPEDEULVNLLPAILSPGALVGVVCAIIRRHVGPBEGA VOMNRLI 1920
QY 1921 AFASRGNHVSPHYVPESDPAARVTQIILSLTITQILKRLHOMINEDCSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPHYVPESDPAARVTQIILSLTITQILKRLHOMINEDCSTPCSGSWLRDV 1980
QY 1981 WDMICTVLDFEKTWLOSKLPLRPGVPLSPGCRGYKGVWRGADGINTQTCPGAQIAGHVK 2040
Db 1981 WDMICTVLDFEKTWLOSKLPLRPGVPLSPGCRGYKGVWRGADGINTQTCPGAQIAGHVK 2040
QY 2041 NGSMLYVGRTPCSNTWHTGFPINATTTGCTSPAPNYSRALAMRYAAEEYEVETRGDGH 2100
Db 2041 NGSMLYVGRTPCSNTWHTGFPINATTTGCTSPAPNYSRALAMRYAAEEYEVETRGDGH 2100
QY 2101 YVTGMTTDNVKKPCPOVPAPEFTEVDGVRLHRYAPACKFLREDEYTFQVGLNOYVLSQI 2160
Db 2101 YVTGMTTDNVKKPCPOVPAPEFTEVDGVRLHRYAPACKFLREDEYTFQVGLNOYVLSQI 2160
QY 2161 PCEPEBDVTVLTSMLTDPESHITAEFTAKRRLARGSPSLAASSASQISAPSLKATCTTHHD 2220
Db 2161 PCEPEBDVTVLTSMLTDPESHITAEFTAKRRLARGSPSLAASSASQISAPSLKATCTTHHD 2220
QY 2221 SPDADILEANLLMRQMGENTRIVESKNVYLLDSFEPILHAGDREISVAAEIIILAKSKK 2280
Db 2221 SPDADILEANLLMRQMGENTRIVESKNVYLLDSFEPILHAGDREISVAAEIIILAKSKK 2280

QY 2281 FFSALPIWAPRPNPPLLESMDPDVVPVVGCCPLPTKAPPIPPRRKRITVLTESNV 2340
 DB 2281 FFSALPIWAPRPNPPLLESMDPDVVPVVGCCPLPTKAPPIPPRRKRITVLTESNV 2340
 QY 2341 SSALIELAKRTGSSSSAVDSGTATALDLDASDDDKSDVESYSMPLEGEPPDPL 2400
 DB 2341 SSALIELAKRTGSSSSAVDSGTATGPPDQASDDDKSDVESYSMPLEGEPPDPL 2400
 QY 2401 SSGSNSTVSEASSEDVCCSMSTWTGALTTPCAEESKLPINPLNSLLRHNNVATY 2460
 DB 2401 SSGSNSTVSEASSEDVCCSMSTWTGALTTPCAEESKLPINPLNSLLRHNNVATY 2460
 QY 2461 SSSASLRQKKVTFEDRLQVLDHYRDVLEKMKAKASTVKAULSIEBCKLTPHSAKSY 2520
 DB 2461 SSSASLRQKKVTFEDRLQVLDHYRDVLEKMKAKASTVKAULSIEBCKLTPHSAKSY 2520
 QY 2521 GYGAKDVNLSRAVNHNSWEDLLEDTETPIDTTIMAKSEVFCQPEKGGKPARLIY 2580
 DB 2521 GYGAKDVNLSRAVNHNSWEDLLEDTETPIDTTIMAKSEVFCQPEKGGKPARLIY 2580
 QY 2581 FPDLPGRVCEKALYVNSTLPQAVMGSSYGQYSPKQVEFLVNTWKSCKCPMGSYDT 2640
 DB 2581 FPDLPGRVCEKALYVNSTLPQAVMGSSYGQYSPKQVEFLVNTWKSCKCPMGSYDT 2640
 QY 2641 RCPDSTVTESDIVBESSIVCCDLAPARQAIRSLTERLYIGGPLTNSKQNGYRCRA 2700
 DB 2641 RCPDSTVTESDIVBESSIVCCDLAPARQAIRSLTERLYIGGPLTNSKQNGYRCRA 2700
 QY 2701 SGVLTSSCGNTLCYLKATAACRAAKLQDCTMLVNGDDLVIIESAGTOBDAALRAFTE 2760
 DB 2701 SGVLTSSCGNTLCYLKATAACRAAKLQDCTMLVNGDDLVIIESAGTOBDAALRAFTE 2760
 QY 2761 AMTRYAPSPDPPQPEYDLLELITSCSSNVSAVDASGRVYVLTTRDPTPLARAMEYAR 2820
 DB 2761 AMTRYAPSPDPPQPEYDLLELITSCSSNVSAVDASGRVYVLTTRDPTPLARAMEYAR 2820
 QY 2821 HTDINSMLGNIIMVAPLTMARMLIMTFESLILAOEALRDLCOITGACYSIEPLDLPQ 2880
 DB 2821 HTDINSMLGNIIMVAPLTMARMLIMTFESLILAOEALRDLCOITGACYSIEPLDLPQ 2880
 QY 2881 IIRERHLSAFTLHSHSGEININRASCRLKLGVPPLTWHRBASVAKLISGGRAATC 2940
 DB 2881 IIRERHLSAFTLHSHSGEININRASCRLKLGVPPLTWHRBASVAKLISGGRAATC 2940
 QY 2941 GRVLFNNAVATKLTLPAPASQDLDSGMFVAGYSGGDIYHSLSRAPRWEPLCLILLSV 3000
 DB 2941 GRVLFNNAVATKLTLPAPASQDLDSGMFVAGYSGGDIYHSLSRAPRWEPLCLILLSV 3000
 QY 3001 GVGITYLLPNR 3010
 DB 3001 GVGITYLLPNR 3010

RESULT 2
 GNAVTC
 genome polyp/Protein - hepatitis C virus
 N:Contains: capsid protein C, envelope protein M, hepacivirin (EC 3.4.21.98) (nonstructu
 protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
 C:Species: hepatitis C virus
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
 C:Accession: A38465
 R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;
 J. Virol. 65, 1105-1113, 1991
 A:Title: Structure and organization of the hepatitis C virus genome isolated from human
 A:Reference number: A38465; MUID:91140698; PMID:1847440
 A:Accession: A38465
 A:Molecule type: genomic RNA
 A:Residues: 1-3010 <TRK>
 A:Cross-references: UNIPROT:P26663; UNIPARC:UPI0000131E1C; EMBL:M58335; NID:g329770; PID
 C:Superfamily: hepatitis C virus genome polyp/Protein
 C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
 F:2-115/Product: capsid protein C #status predicted <CPC>
 F:116-191/Product: envelope protein M #status predicted <ERM>

F:192-389/Product: major envelope protein E #status predicted <ME>
 F:190-729/Product: nonstructural protein NS1 #status predicted <NS1>
 F:130-1006/Product: nonstructural protein NS2 #status predicted <NS2>
 F:1007-1615/Product: hepacivirin #status predicted <NS3>
 F:1230-1337/Region: nucleotide-binding motif A (P-loop)
 F:1312-1317/Region: nucleotide-binding motif B
 F:1316-1319/Region: DEXH motif
 F:1616-1662/Product: nonstructural protein NS4a #status predicted <NS4>
 F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4>
 F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
 F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match 95.3%; Score 15253; DB 1; Length 3010;
 Best Local Similarity 94.4%; Pred. No. 0;
 Matches 2840; Conservative 78; Mismatches 92; Indels 0; Gaps 0;

QY 1 MSTNPKRQKTKRNTRRPODYKFPGGQIVGVVLLPRRGRPLGYRATRKASERSQPRG 60
 DB 1 MSTNPKRQKTKRNTRRPODYKFPGGQIVGVVLLPRRGRPLGYRATRKASERSQPRG 60
 QY 61 RROPIDKARRPEGRAWOPGYPMPLVNGELGNAWMLSPRGRSPSGPTDPRRRSRLNG 120
 DB 61 RROPIDKARRPEGRAWOPGYPMPLVNGELGNAWMLSPRGRSPSGPTDPRRRSRLNG 120
 QY 121 KVIDLITCGFADIMGYIPLVGAFLGGAARALANGVLEGVNVAATGNLPGCSFSTFLA 180
 DB 121 KVIDLITCGFADIMGYIPLVGAFLGGAARALANGVLEGVNVAATGNLPGCSFSTFLA 180
 QY 181 LLSCLTTPASAYEVNRSVSGTYHTNDNSNSIYEEADVIMHPRGCVPCVQEGNSSRCW 240
 DB 181 LLSCLTTPASAYEVNRSVSGTYHTNDNSNSIYEEADVIMHPRGCVPCVQEGNSSRCW 240
 QY 241 ALPTLAARVASPTTIRHVDLAVGTAFCSAMVYGDLCGSIPLVSQLTFSPPRHET 300
 DB 241 ALPTLAARVASPTTIRHVDLAVGTAFCSAMVYGDLCGSIPLVSQLTFSPPRHET 300
 QY 301 VODNCSIIYGHVSGHRMAMDMNMSPTTALVVSQILRIPOAVVDVAVAGHMGVLAGLA 360
 DB 301 VODNCSIIYGHVSGHRMAMDMNMSPTTALVVSQILRIPOAVVDVAVAGHMGVLAGLA 360
 QY 361 YSSVNGMAKVLIVALLFAGVDGETHTTGRVAGHTTSGLTSPSSGASOKIQLVNTNGSW 420
 DB 361 YSSVNGMAKVLIVALLFAGVDGETHTTGRVAGHTTSGLTSPSSGASOKIQLVNTNGSW 420
 QY 421 HINRTALNCNDSLOTGFPAALFPAHKFNSSGCCPRMASCPIDMFAQMGPIITTKRNS 480
 DB 421 HINRTALNCNDSLOTGFPAALFPAHKFNSSGCCPRMASCPIDMFAQMGPIITTKRNS 480
 QY 481 DQRPYCHYAPRPGGVPAQOVGCPVCTPSPVNVGTTDRSGVPTVSMGENETDWMLN 540
 DB 481 DQRPYCHYAPRPGGVPAQOVGCPVCTPSPVNVGTTDRSGVPTVSMGENETDWMLN 540
 QY 541 NTRPPGNNWFGCTMNSGTGFTKCGPPCNIIGVGNRTLICPTDCFRKHPEATYTKCSG 600
 DB 541 NTRPPGNNWFGCTMNSGTGFTKCGPPCNIIGVGNRTLICPTDCFRKHPEATYTKCSG 600
 QY 601 PMLPRCIVDYPRYLMHYPTCLNFSIRKRVNYVGVGEHRLNAACNMTGERCNLEDRDS 660
 DB 601 PMLPRCIVDYPRYLMHYPTCLNFSIRKRVNYVGVGEHRLNAACNMTGERCNLEDRDS 660
 QY 661 ELSPILLSTTEMQILPCAFPTTLPALSTGLIHLDONIVDVQILYGVGSAFVSFAIKMEYL 720
 DB 661 ELSPILLSTTEMQILPCAFPTTLPALSTGLIHLDONIVDVQILYGVGSAFVSFAIKMEYL 720
 QY 721 LFLFLADARVACACIMMMLTIAQBALENIIVLNASVAGAHGILSLVFCFCAMYIKG 780
 DB 721 LFLFLADARVACACIMMMLTIAQBALENIIVLNASVAGAHGILSLVFCFCAMYIKG 780
 QY 781 RLAPQAAVAFYGVNPLILLALPPRAYALDRMAASCGAVLVGLVFLTSPYKVFYT 840
 DB 781 RLAPQAAVAFYGVNPLILLALPPRAYALDRMAASCGAVLVGLVFLTSPYKVFYT 840
 QY 841 RLIMWLQYFTRAEAHQVWVPLNVNRRGDAIILITCAVHPELIIDITKLILALIGPLM 900

RESULT 3

A45573 genome polyprotein - hepatitis C virus (strain J7)

N:contains: capsid protein C; envelope protein M; hepatitis C virus (HCV) genome from a single Japanese carrier: S

protein NS4; nonstructural protein NS4b; nonstructural protein NS5

C:species: hepatitis C virus

C:update: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004

C:accession: A45573

R:rank: T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,

Virus Res. 23, 39-53, 1992

A:title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: S

A:reference number: A45573; PMID:9225714; PMID:1318627

A:accession: A45573

A:status: preliminary

A:molecule type: DNA

A:residues: 1-3010 <TAN>

A:cross-references: UNIPROT:000269; UNIPARC:UP10000131E29; GB:D11168; GB:D01171; NID:922

A:experimental source: HCV-JT

A:note: sequence extracted from NCBI backbone (NCBI:106206, NCBI:106207)

C:superfamily: hepatitis C virus genome polyprotein

C:keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin

F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <EPM>

F:192-389/Product: major envelope protein E #status predicted <ME>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F:1007-1615/Product: hepatitis C virus status predicted <NS3>

F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B

F:1316-1319/Region: DEX motif

F:1616-1862/Product: nonstructural protein NS4 #status predicted <NS4>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 95.2%; Score 15236; DB 1; Length 3010;

Best Local Similarity 94.1%; Pred. No. 0;

Matches 2833; Conservative 83; Mismatches 94; Indels 0; Gaps 0;

QY 1 MSTNPKQRKTKRTNRPOVKPPGGQIVGVYLLPRRGRPLGVATRKASRSQPRG 60
DB 1 MSTNPKQRKTKRTNRPOVKPPGGQIVGVYLLPRRGRPLGVATRKASRSQPRG 60
QY 61 RROPIPRARREGRAMAOPGYWPLYGNEGLGMAWMLSPRGSPPSGPTDPRRSRLG 120
DB 61 RROPIPRARREGRAMAOPGYWPLYGNEGLGMAWMLSPRGSPPSGPTDPRRSRLG 120
QY 121 KVIDITLTCGFPADLMGYIPLVGAPLGGARALAHGVRVLEDDVNAATGNLPCCSFSIFLLA 180
DB 121 KVIDITLTCGFPADLMGYIPLVGAPLGGARALAHGVRVLEDDVNAATGNLPCCSFSIFLLA 180
QY 181 LLSCLTIPASAYEVNRSYGIHVTNDCNSNIIYEAADVIMHTPGCVCEGNSRRCMV 240
DB 181 LLSCLTIPASAYEVNRSYGIHVTNDCNSNIIYEAADVIMHTPGCVCEGNSRRCMV 240
QY 241 ALPTTLAARNASVPTTIRRHVDLLVGTAFCSAAYVVDLCSIFLVLSQFTFSRRHET 300
DB 241 ALPTTLAARNASVPTTIRRHVDLLVGTAFCSAAYVVDLCSIFLVLSQFTFSRRHET 300
QY 301 VQDCNCSYIPGHVSGHRAAMMMNSPTTALVVSQLLRIPOAVVDMAVGAHMGVLADLA 360
DB 301 VQDCNCSYIPGHVSGHRAAMMMNSPTTALVVSQLLRIPOAVVDMAVGAHMGVLADLA 360
QY 361 YYSWVGNMAKVLVALLPAGVDGETHTTGVAGHTTSGFTSLFSSGASOKIQLVNTNSW 420
DB 361 YYSWVGNMAKVLVALLPAGVDGETHTTGVAGHTTSGFTSLFSSGASOKIQLVNTNSW 420
QY 421 HINRTALNCNDSLOTGFPALFYAHKFNSSGCPERMASSCPIDMFAQCGMPTTYTKPNSS 480
DB 421 HINRTALNCNDSLOTGFPALFYAHKFNSSGCPERMASSCPIDMFAQCGMPTTYTKPNSS 480
QY 481 DORPYCMHYARPCGVVPAQVCGVYCTPSPVTVVGTDRSGVFTYSMGNEEDVMLLN 540

DB 481 DORPYCMHYARPCGVVPAQVCGVYCTPSPVTVVGTDRSGVFTYSMGNEEDVMLLN 540
QY 541 NTRPPGNNFGCTWNSSTGFTKTCGGPPCNIGGVNRTILCPDFCRKIPKATYTCGSG 600
DB 541 NTRPPGNNFGCTWNSSTGFTKTCGGPPCNIGGVNRTILCPDFCRKIPKATYTCGSG 600
QY 601 PMLTPRCLVDYPRMLHYCTVNFITFKRYMYGVGEHLSAACNMTGRRCNLEDRDS 660
DB 601 PMLTPRCLVDYPRMLHYCTVNFITFKRYMYGVGEHLSAACNMTGRRCNLEDRDS 660
QY 661 ELSPLLSTTEWQILPCATTLPALSTGLIHHQNVVDVQYLYGVCSAFVSPAKMEYIL 720
DB 661 ELSPLLSTTEWQILPCATTLPALSTGLIHHQNVVDVQYLYGVCSAFVSPAKMEYIL 720
QY 721 LFLLLADARVCCIMMMLLIQAEEALELVLANAASLAGADGILSPVFCAMYIKG 780
DB 721 LFLLLADARVCCIMMMLLIQAEEALELVLANAASLAGADGILSPVFCAMYIKG 780
QY 781 RLAPGAAYFVGVPLLLLLALPPRAYALDREMAASCGAVLVGLVFTLSPYKVELT 840
DB 781 RLAPGAAYFVGVPLLLLLALPPRAYALDREMAASCGAVLVGLVFTLSPYKVELT 840
QY 841 RLIMWLQYFTRAEAHMQVWVPLNVRGGRDAIILLTCAVHPELIDITKLLAILGRLM 900
DB 841 RLIMWLQYFTRAEAHMQVWVPLNVRGGRDAIILLTCAVHPELIDITKLLAILGRLM 900
QY 901 VLOAGITRVYFPRAQGLIACMLVRKVGAGHVQVWFMKLGALTGVYVNHITLPLDMA 960
DB 901 VLOAGITRVYFPRAQGLIACMLVRKVGAGHVQVWFMKLGALTGVYVNHITLPLDMA 960
QY 961 HAGRLDAVAVEPVVFSAMETKVIITWGAADTAACGDIILGHPVARSKEIFLGPADSLG 1020
DB 961 HAGRLDAVAVEPVVFSAMETKVIITWGAADTAACGDIILGHPVARSKEIFLGPADSLG 1020
QY 1021 QGWRLLAPITVAYSQOTRGVLGCIITSLTGRDKQVBEQVQVSTATQSLATCINGVCM 1080
DB 1021 QGWRLLAPITVAYSQOTRGVLGCIITSLTGRDKQVBEQVQVSTATQSLATCINGVCM 1080
QY 1081 VYHAGSKTLAGKGPITQWYTVNVDLVGQAPRARSMTPGCGSSDLYLVRHADVI 1140
DB 1081 VYHAGSKTLAGKGPITQWYTVNVDLVGQAPRARSMTPGCGSSDLYLVRHADVI 1140
QY 1141 PVRRRGDSGLSPRVSVLYKSSGGPILCPGSHVGVFRAACTGSAKAVDFIPVES 1200
DB 1141 PVRRRGDSGLSPRVSVLYKSSGGPILCPGSHVGVFRAACTGSAKAVDFIPVES 1200
QY 1201 METTMRSPVFTDNSTPRAVQTFQVAILHAPTSGSKSTKVPAAYAAQGYKVLVNPVAA 1260
DB 1201 METTMRSPVFTDNSTPRAVQTFQVAILHAPTSGSKSTKVPAAYAAQGYKVLVNPVAA 1260
QY 1261 TLFGAYMSKAHGIIDPNIIRGVRTITTTGSGITSTYTKFLADGCGSGADYIICDECHS 1320
DB 1261 TLFGAYMSKAHGIIDPNIIRGVRTITTTGSGITSTYTKFLADGCGSGADYIICDECHS 1320
QY 1321 TDSFTTIIIGIGVLDQAEATAGARLVVLAATAPRGSVTPHPNIEEISNNGEIPFYKAI 1380
DB 1321 TDSFTTIIIGIGVLDQAEATAGARLVVLAATAPRGSVTPHPNIEEISNNGEIPFYKAI 1380
QY 1381 PLEAIKGGRHILFCHSKKCDLAAKLSGLGINNAVYRGIDVGIPIPSGVVIVATDAL 1440
DB 1381 PLEAIKGGRHILFCHSKKCDLAAKLSGLGINNAVYRGIDVGIPIPSGVVIVATDAL 1440
QY 1441 MTGFTGDPDSVIDNCTVOTVDPFSLDPTFTIETTVYPODAVNSORRGRTGRGSGIYR 1500
DB 1441 MTGFTGDPDSVIDNCTVOTVDPFSLDPTFTIETTVYPODAVNSORRGRTGRGSGIYR 1500
QY 1501 FVTGERSGMPDSSVCEQDAGACAWELPRAETSVLRAYLMTGPIPVQODHLEFMS 1560
DB 1501 FVTGERSGMPDSSVCEQDAGACAWELPRAETSVLRAYLMTGPIPVQODHLEFMS 1560
QY 1561 VFTGLTIDAHFLSOTKQAGDNFYLVAVQAATVCARAAPPSSDQWCKLIRLPTLHG 1620
DB 1561 VFTGLTIDAHFLSOTKQAGDNFYLVAVQAATVCARAAPPSSDQWCKLIRLPTLHG 1620

```
Db 1561 VFTGLTHIDAHFLSOTKQAGDNFVVAQAQVCAQAQAPPSNDQMMKCLIRKPTLHG 1620
Qy 1621 PFPLLYRLGAVONEVILTHPTKYIMACNSADLEVVTSWLVGVLAALAAVCLTTGSGV 1680
Db 1621 PFPLLYRLGAVONEVILTHPTKYIMACNSADLEVVTSWLVGVLAALAAVCLTTGSGV 1680
Qy 1681 VVGVTIILSGKAVVDPREVLYOEPRMEECASOLPYTRCQMOAQQFOKALGLLQTX 1740
Db 1681 VVGVTIILSGKAVVDPREVLYOEPRMEECASOLPYTRCQMOAQQFOKALGLLQTX 1740
Qy 1741 KQAEAAAPVESKMRALFTFMKAMNFIISGIOYLAGLSTLPQNPALIASIMAFASITSP 1800
Db 1741 KQAEAAAPVESKMRALFTFMKAMNFIISGIOYLAGLSTLPQNPALIASIMAFASITSP 1800
Qy 1801 LTTQNTLLFNILGGVAAQALAPPSAASAFVAGIAGAAVSGIGLKVLYVLIAGYAGVA 1860
Db 1801 LTTQNTLLFNILGGVAAQALAPPSAASAFVAGIAGAAVSGIGLKVLYVLIAGYAGVA 1860
Qy 1861 GALVAFKVMGSEVPSTEDVLLPALLSGALVGVVCAALLRHVGPBGAVQMMNRLI 1920
Db 1861 GALVAFKVMGSEVPSTEDVLLPALLSGALVGVVCAALLRHVGPBGAVQMMNRLI 1920
Qy 1921 AFASRGNHVSPTHYVBPESDAAAVTQILSLITTLQKLHQMINEDCSTPCSGSWLRDY 1980
Db 1921 AFASRGNHVSPTHYVBPESDAAAVTQILSLITTLQKLHQMINEDCSTPCSGSWLRDY 1980
Qy 1981 WDMICTVLDTDFKTWILQSKLLPRLPGVPLSCORGYKVGWRGDIMQTTCPGQAQIAGHV 2040
Db 1981 WDMICTVLDTDFKTWILQSKLLPRLPGVPLSCORGYKVGWRGDIMQTTCPGQAQIAGHV 2040
Qy 2041 NSMRIVGRTCSNTHGTPINAYTTGCTPSPAVNSGALMRVAEEVETRRGDH 2100
Db 2041 NSMRIVGRTCSNTHGTPINAYTTGCTPSPAVNSGALMRVAEEVETRRGDH 2100
Qy 2101 YVTGMTTDVVKCEQVPAPEFTEVDVGLHRYPACKPLREDYFOYGLNOYLVGSOL 2160
Db 2101 YVTGMTTDVVKCEQVPAPEFTEVDVGLHRYPACKPLREDYFOYGLNOYLVGSOL 2160
Qy 2161 PCEPEBDVTLTSMULTDPSHITAEYAKRLIARGSPSLASSSASQLSAPSLKATCTTHD 2220
Db 2161 PCEPEBDVTLTSMULTDPSHITAEYAKRLIARGSPSLASSSASQLSAPSLKATCTTHD 2220
Qy 2221 SPDAULIEANILMROMGNGNITRVESENVVILDSFEPLHAEDDEEISVAAEILKSKK 2280
Db 2221 SPDAULIEANILMROMGNGNITRVESENVVILDSFEPLHAEDDEEISVAAEILKSKK 2280
Qy 2281 PPSALPIYMPDPNPPLLESMDKDPVVPVHGCPPLPTKAPPIPPRRKRTVVLRESNV 2340
Db 2281 PPSALPIYMPDPNPPLLESMDKDPVVPVHGCPPLPTKAPPIPPRRKRTVVLRESNV 2340
Qy 2341 SSALAEALATKTFGSSGSSAVDSGTATAPDLASDDDKSGDVESYSMPPLBEGEPDPL 2400
Db 2341 SSALAEALATKTFGSSGSSAVDSGTATAPDLASDDDKSGDVESYSMPPLBEGEPDPL 2400
Qy 2401 SSGSMSTVSEBESADVCCSMSTTWTGALITPCAABESKLPINPLNSILRHHNMYATT 2460
Db 2401 SSGSMSTVSEBESADVCCSMSTTWTGALITPCAABESKLPINPLNSILRHHNMYATT 2460
Qy 2461 SPSASLROKKVPPDRLOVLDHRYDLKEMKAKASTYKAKLSIEEACKLTPHSAKSKF 2520
Db 2461 SPSASLROKKVPPDRLOVLDHRYDLKEMKAKASTYKAKLSIEEACKLTPHSAKSKF 2520
Qy 2521 GYGAKDVNLSSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVFCVOPKGRKPARLLY 2580
Db 2521 GYGAKDVNLSSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVFCVOPKGRKPARLLY 2580
Qy 2581 FPDLAGRVCEKALYVNSTLPAVNGSSYGFQYSPKORVEFLVNTWKSCKCMGSSYDT 2640
Db 2581 FPDLAGRVCEKALYVNSTLPAVNGSSYGFQYSPKORVEFLVNTWKSCKCMGSSYDT 2640
Qy 2641 RCDSDSVTSDDIRVESIYQCCDLPAPBARQATSLRERLYTGGLPNSKQNGCYRRCRA 2700
Db 2641 RCDSDSVTSDDIRVESIYQCCDLPAPBARQATSLRERLYTGGLPNSKQNGCYRRCRA 2700
```

```
Qy 2701 SGVLTTSQGNLTTCYLKATAACRAKLODCTMLVNGDDLVIYCESAGTQEDAAALRPFTE 2760
Db 2701 SGVLTTSQGNLTTCYLKATAACRAKLODCTMLVNGDDLVIYCESAGTQEDAAALRPFTE 2760
Qy 2761 AMTRYSAAPGDPPEPEYLELITSCSNVSVYHADSGRVYVLTDPPTPLRAAMETAR 2820
Db 2761 AMTRYSAAPGDPPEPEYLELITSCSNVSVYHADSGRVYVLTDPPTPLRAAMETAR 2820
Qy 2821 HTPINSLGNITMVAPTLMARNIMLTHFFSILLAOEOLKALDCQIYGACYSIEPLDPEQ 2880
Db 2821 HTPINSLGNITMVAPTLMARNIMLTHFFSILLAOEOLKALDCQIYGACYSIEPLDPEQ 2880
Qy 2881 IIERLHGSAFTLHSGEINRVASCIRKGVPLRLTWRRPARSVRAKLSQGGRAATC 2940
Db 2881 IIERLHGSAFTLHSGEINRVASCIRKGVPLRLTWRRPARSVRAKLSQGGRAATC 2940
Qy 2941 GRYLPMVAVRTKLKLTPIPAASQDLSCGFVAGYSGDIIYHLSRAPRPFCLLLLSV 3000
Db 2941 GRYLPMVAVRTKLKLTPIPAASQDLSCGFVAGYSGDIIYHLSRAPRPFCLLLLSV 3000
Qy 3001 GVGIVLPLNR 3010
Db 3001 GVGIVLPLNR 3010

RESULT 4
GNMVTW
genome polypeptide - hepatitis C virus (strain Taiwan)
N:contains: capsid protein C; envelope protein M; hepacivirin (BC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:species: hepatitis C virus
A:note: host Homo sapiens (man)
C:date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:accession: A40244
R:chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A:title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A:reference number: A40244; MUID:92230206; PMID:1314449
A:molecule type: genomic RNA
A:residues: 1-3010 <CH>
A:cross-references: UNIPROT:P29846; UNIPARC:UPI0000131E2B; GB:M84754
C:superfamily: hepatitis C virus genome polypeptide
C:keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2071.

Query Match 94.1%; Score 15062; DB 1; Length 3010;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2807; Conservative 81; Mismatches 122; Indels 0; Gaps 0;

Qy 1 MSTNPKQKRTKRNTRRRPODVKFGGQIYGVVLLPRGRLGVRATRKASERSQPRG 60
Db 1 MSTNPKQKRTKRNTRRRPODVKFGGQIYGVVLLPRGRLGVRATRKATKWSQPRG 60

Qy 61 RQQIPKARPPGRAMAOPGYWPPLYGNEGICAGMLSPGSRPSWGPTDPRRSRNIG 120
Db 61 RQQIPKARPPGRAMAOPGYWPPLYGNEGICAGMLSPGSRPSWGPTDPRRSRNIG 120

Qy 121 KYIDTLTGCFALMGYIPIVGAALGGAARALAHGVVLEDGVNVTAGNLPGCSFSFLA 180
Db 121 KYIDTLTGCFALMGYIPIVGAALGGAARALAHGVVLEDGVNVTAGNLPGCSFSFLA 180
```

Db 121 KVDTLTCGFADLMGYIPLVGAFLGGVARALAHGVRLEDCGVNATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVYEADVTIMHTPGCVPCVQEGNSSRCWV 240
Db 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVYEADVTIMHTPGCVPCVRENSSRCWV 240
Qy 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVGDLCGSIPLVSQFLFTSPRRHET 300
Db 241 ALPTTLAARNNSVPTAIRRHVDLLVGAFAFCSAMVGDLCGSVFLVSQFLFTSPRRYET 300
Qy 301 VQDNCNSIYPGHVSGHRMADMNNSPTTALVVSQLLRIPQAVVDMVACAHGVGLAGLA 360
Db 301 VQDNCNSIYPGHVSGHRMADMNNSPTTALVVSQLLRIPQAVVDMVACAHGVGLAGLA 360
Qy 361 YYSNVGNWAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKQLVNTNGSW 420
Db 361 YYSNVGNWAKVLIVMLLFAGVDGSTIVSGGTVARTTHSLASLFTQOGASQIKQLINTNGSW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCCPERMAASCRPIDWFAOGWGPITYTKPNSS 480
Db 421 HINRTALNCNDSLOTGFLASLIFYAHREFNAGCCPERMAASCRSIDKFDQGWGPITYTEADIQ 480
Qy 481 DORPYCHHYAPRPCGVVPASQVCPVYCFTPSPVVVGTDRSGVPTYSGENETDVMLLN 540
Db 481 DORPYCHHYAPRPCGVVPASQVCPVYCFTPSPVVVGTDRFGAPTYSGENETDVLILN 540
Qy 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPNCNIGVGNRTLICPTDCFRKHPPEATYKCGSG 600
Db 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPNCNIGGGNNTLVCPTDCFRKHPPEATYKCGSG 600
Qy 601 PMLTPRCLVDYPYRLMHYPCTLNPSIFKVRMYGVGVEHRLNAACNWTRGERCNLEDRDRS 660
Db 601 PMLTPRCMVDPYRLMHYPCVNTFIFKVRMYGVGVEHRLNAACNWTRGERCNLEDRDRS 660
Qy 720 ELSPLLLSTTEWQILPCAFTLLPALSTCLLHLHQNIVDVQVLYCGVSFAFSAFKWEYL 720
Db 661 ELSPLLLSTTEWQILPCSFTGLPALSTGLLHLHQNVDVQVLYGIGSAVVSFAFKWEYL 720
Qy 721 LFLFLLADARVACACLWMMLLTAQEAALENILVNLNAAVAGAHGILSPLVFFCAAWYIKG 780
Db 721 LFLFLLADARVACACLWMMLLTAQEAALENILVFNAAVAGWHGTLSPFLVFFCAAWYIKG 780
Qy 781 RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCGGAVVGLVPLTLPSPYKVFPLT 840
Db 781 RLVPGAAYALYGVWPLLLLLALPPRAYALDREMAASCGGAVVGLVPLTLPSPYKWFPLA 840
Qy 841 RL1WLWQYFTRAEAHQWVPLNVGRGRDAIILLTCVNHPELIFDITKLLAILGLPM 900
Db 841 RL1WLWQYFTRAEAHQWVPLNVGRGRDAIILLTCVNHPELIFDITKLLAILGLPM 900
Qy 901 VLQAGITRPVYFVRAOGLIRACMLVRKVAGGHYQVMFPMKALGTGVVYNHLLTPLRDWA 960
Db 901 VLQAGLTRIPYFVRAOGLIRACMLVRKAAGHYVQMALMKUALITGVVYDHLTPLQDWA 960
Qy 961 HAGRLDVAVEPVVFSAMETKVITWGDATTAACGDIILGLPVSARRKEIFLGPADSLEG 1020
Db 961 HTGLRDLVAVEPVVFSDMETKIIWGDATTAACGDIILGLPVSARRGREILLGPADSLEG 1020
Qy 1021 QGWRLLAPITAYSOQTRGVLCIITSLTGRDNQVGEVQVVSSTATQFLATCINGVCWT 1080
Db 1021 RQWRLLAPITAYAOQTRGLFGCIITSLTGRDNQVGEVQVVSSTATQFLATCINGVCWT 1080
Qy 1081 VYHGAGSKTLAGPKGPIQMTYNTVDLDIVGWAOPGARSMTPCSCGSSDLYLVTRHADVI 1140
Db 1081 VYHGAGSKTLAGPKGPIQMTYNTVDQDLVGHAPQGARSLTPTCTCGSSDLYLVTRHADVI 1140
Qy 1141 PVRRGDSRGILLSRPRYSYLKSGSGPGLLCPSGHVGVVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRGDSRGILLSRPRISYLYKSGSGPGLLCPSGHVGVVFRAAVCTRGVAKAVDFIPVES 1200
Qy 1201 METTMRSPVFTDNTSPAPVQFQVLAHLHAPTGSCKSTKVPAAVAAQYKVLVLPNSVAA 1260
Db 1201 METTMRSPVFTDNTSSPPVQAFQVLAHLHAPTGSCKSTKVPAAVAAQYKVLVLPNSVAA 1260

Qy 1261 TLGFGAYMSKAHGDIPNIRTGVRTITTCGSIITYSTYKFLADGCCSGGAYDIILCDECHS 1320
Db 1261 TLGFGAYMSKAHGVDPNIRTGVRTITTCGSIITYSTYKFLADGCCSGGAYDIILCDECHS 1320
Qy 1321 TOSTTTLIGTIVLDOAETAGARLVLATATPPGCVTPVPHNIBEEIGLSNNGEIPFYFKGAI 1380
Db 1321 TOSTTTLIGTIVLDOAETAGARLVLATATPPGCVTPVPHNIBEEIGLSNNGEIPFYFKGAI 1380
Qy 1381 PIEAIKGRHILFCHSKKKKDELAALKLGLGLNNAVAYYRGLDVSVPIPIGVVVVATDAL 1440
Db 1381 PIETIKGRHILFCHSKKKKDELAALKLSALGIHNAVAYYRGLDVSVIPASGNVVVATDAL 1440
Qy 1441 MTGFTGDFSVIDCNTCVTQTVDFSLDPTFTTETTTVPDAYSRSORRGRTGRSGIYR 1500
Db 1441 MTGFTGDFSVIDCNTCVTQTVDFSLDPTFTTETTTVPDAYSRSORRGRTGRSGIYR 1500
Qy 1501 FVTPGGRPSGMFSSVLCYDAGCAWYELTPAETSVRRLRAYLNTPLGVPQCDBLEFWES 1560
Db 1501 FVTPGGRPSGMFSSVLCYDAGCAWYELTPAETSVRRLRAYLNTPLGVPQCDBLEFWES 1560
Qy 1561 VFTGLTHIDAHFLSOTKQAGDNPPYLVAOATVCARAQAPPPSWDQMWKCLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSOTKQAGDNPPYLVAOATVCARAQAPPPSWDQMWKCLIRLKPTLHG 1620
Qy 1621 PTPLLYRLGAVQNEVILTHPIITKYIMACHSADLEVVTSTWLVGGVLAALAAAYCLTTGVS 1680
Db 1621 PTPLLYRLGAVQNEVILTHPIITKYIMACHSADLEVVTSTWLVGGVLAALAAAYCLTTGVS 1680
Qy 1681 VIYGRILILSKPAVPPDREVLVYQEFDEMEECASQOLPYIEQGMOLAEQFKKALGLLQAT 1740
Db 1681 VIYGRILILSKPAVPPDREVLVYQEFDEMEECASQOLPYIEQGMOLAEQFKKALGLLQAT 1740
Qy 1741 KQAEAAAPVVESKWRALETFMAKHMNFISIGIOYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KQAEAAAPVVESKWRILEAFWANDMNFISIGIOYLAGLSTLPGNPAIASLMAFTASITSP 1800
Qy 1801 LTTQNTLLFNIILGWWAAQLAPSAASAFVAGIAGAAVCSIGLGKVLVDILAGYGGAVA 1860
Db 1801 LTTQNTLLFNIILGWWAAQLAPSAASAFVAGIAGAAVCSIGLGKVLVDILAGYGGAVA 1860
Qy 1861 GALVAFKVMGSEVPSTEDIVNLLPALILSPCALVGVVCAAILRRHVGPGEVAGVOMNRLI 1920
Db 1861 GALVAFKVMGSEVPSTEDIVNLLPALILSPCALVGVVCAAILRRHVDPGEVAGVOMNRLI 1920
Qy 1921 AFASRGNHVSPHYVPESDAAARVTQILSLTITQLLKRHLQWINEDCSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPHYVPESDAAARVTQILSLTITQLLKRHLQWINEDCSTPCSGSWLRDV 1980
Qy 1981 WDMICTVLTDFTKWLOSKLLPRLPGVPPFLSCQRYKGVWRGDMQITTCPCGAQIAGHYK 2040
Db 1981 WDMICTVLADFTKWLOSKLLPRLPGVPPFLSCQRYKGVWRGDMQITTCPCGAQITGHVK 2040
Qy 2041 NGSMRIYGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALWVAEEVVEVTRVGDFH 2100
Db 2041 NGSMRIYGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALWVAEEVVEVTRVGDFH 2100
Qy 2101 YVTGMTDNNKPCQVCPAPPEFFTEVDGVRLHRYAPACKPLLRDVTTFQVGLNQYLVGSQL 2160
Db 2101 YVTGMTDNNKPCQVCPAPPEFFTEVDGVRLHRYAPACKPLLRDVTTFQVGLNQYLVGSQL 2160
Qy 2161 PCEPEPDVTLTSMLTDPSSHITAEAKRRLARGSPPLSASSASQLSAPSLKACTTTHHD 2220
Db 2161 PCEPEPDVTLTSMLTDPSSHITAEAKRRLARGSPPLSASSASQLSAPSLKACTTTHHD 2220
Qy 2221 SPDADLLEANLLWQEMGNNITRVESENKVILDSFPLHAEGDEREISVAEILRKSRK 2280
Db 2221 PPDADLLEANLLWQEMGNNITRVESENKVILDSFPLHAEGDEREISVAEILRKSRK 2280
Qy 2281 FPSALPTWARPDYNNPPLLESWKDPDYVPVVGCPPLPPTKAPPIPPPRKKTIVLTESTV 2340
Db 2281 FPPALPVWARPDPYNNPPLLESWKDPDYVPVVGCPPLPPTKAPPIPPPRKKTIVLTESTV 2340

QY	2341	SSALAEATKTRGSSGSSAVDGTATAPDLASDDGDKGSDVESYSSMPPLGEDEPDPL	2400
Db	2341	SSALAEATKTRGSSGSSAGGTAATAPDPQSDGDDAGSDVESCSMPPLGEDEPDPL	2400
QY	2401	SDGSNTVSEESASDVVCCSMSTWTGALITTCABESKLPINPLNSLLRHHNMVYATT	2460
Db	2401	SDGSNTVSEEBEGEVICCSMSYTWGALITTCABESKLPINALNSLLRHHNMVYATT	2460
QY	2461	SRSASLRQKVTFRDLQVLDDHYRDVLEKMKAKASTVKAKLLSIEBACKLTPPHSAKSF	2520
Db	2461	SRSASORQKVTIDRLQVLDDHYRDVLEKMKAKASTVKAKLLSIEBACKLTPPHSARSKF	2520
QY	2521	GYGAKDVRNLSRAVNHRSVMEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV	2580
Db	2521	GYGAKDVRNLSGKAINHRSVMKDLLETETPIDTTIMAKNEVFCVQPEKGRKPARLIV	2580
QY	2581	FPDLGVRVCEKALYDVVSTLPOAVMGSSYGFGQYSPQKRVFLVNTWKKCPMGFSYDT	2640
Db	2581	FPDLGVRVCEKALYDVVSTLPOAVMGSSYGFGQYSPQKRVFLVNAWKKCPMGFSYDT	2640
QY	2641	RCFDSVTSTSDIRVEESIYCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGYRRCRA	2700
Db	2641	RCFDSVTSTSDIRVEESIYCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGYRRCRA	2700
QY	2701	SGVLTTSCTGNTLCYLKATAACRAAKLQDCTMLVNGDDLJVICESAGTQEDAAALRAFTF	2760
Db	2701	SGVLTTSCTGNTLCYLKASACRAAKLQDCTMLVCGDDLJVICESAGTQEDAAALRVFTF	2760
QY	2761	AMTRYSAPEGDPQPPDYDELITSCSSNVSVAHDSAGKRVYVLTDRPTPLARAAWETAR	2820
Db	2761	AMTRYSAPEGDLQPPDYDELITSCSSNVSVAHDSAGKRVYVLTDRPTPLARAAWATAR	2820
QY	2821	HTPINSWLGNIIMYAPTLWARMLMTHFFSILLAOEKLKALDCQLYGACYSEIPLDLPO	2880
Db	2821	HTPINSWLGNIIMYAPTLWARMLMTHFFSILLAOEKLKALDCQLYGACYSEIPLDLPO	2880
QY	2881	IIERHLGSAFTLHYSVSGEINRVASCLRKLGVPPLRTWRHRARSVRAKLLSQGGRAATC	2940
Db	2881	IIERHLGSAFSLHSVSGEINRVASCLRKLGVPPLRANHRARSVRAKLLSQGGRAATC	2940
QY	2941	GRYLFNNAVTKLKLTPIPAASOLDLSGMFWFVAGYSGDDIYHLSRARPPWFPLCULLLSV	3000
Db	2941	GRYLFNNAVTKLKLTPIPAASOLDLSKFWFVAGYGGDDIYHLSRARPPWFMLCULLLSV	3000
QY	3001	GVGIYLLPNR 3010	
Db	3001	GVGIYLLPNR 3010	

RESULT 5

S18030
genome polypotein - hepatitis C virus (isolate JK1)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4; nonstructural
C:Species: hepatitis C virus
A:Variety: isolate JK1
C:Date: 19-May-2000 #sequence revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S18030; S33570; A48332; S18029
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: A whole genome of hepatitis C virus cdna was isolated from a single patie
A:Reference number: S18028
A:Accession: S18030
A:Molecule type: genomic RNA
A:Residues: 1-3010 <HOW>
A:Cross-references: UNIPROT:Q68949; UNIPARC:UPI00000F2A81; EMBL:X61596; NID:g59478; PIDN
A:Experimental source: isolate JK1 from an individual
R:Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A:Reference number: A48332; MUID:93119270; PMID:8380322
A:Accession: S33570
A:Molecule type: genomic RNA

A:Residues: 1-547, 'T', '549-621, 'V', '623-624, 'S', '626-652, 'DL', '655-761, 'T', '763-782 <HOW>	
A:Cross-references: UNIPARC:UPI00001749FF; EMBL:X61591	
A>Note: this sequence is inconsistent with the nucleotide translation	
A>Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320 as	
as Trp, and TTC for residue 771 as Ser	
A>Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIPI:121748)	
C:Superfamily: hepatitis C virus genome polyprotein	
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serine	
F:116-191/Product: capsid protein C #status predicted <CPC>	
F:192-389/Product: major envelope protein E #status predicted <EPM>	
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>	
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>	
F:1007-1615/Product: hepacivirin #status predicted <NS3>	
F:1230-1237/Region: nucleotide-binding motif A (P-loop)	
F:1312-1317/Region: nucleotide-binding motif B	
F:1316-1319/Region: DEXH motif	
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>	
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>	
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>	
F:196, 209, 234, 250, 305, 417, 423, 448, 532, 540, 556, 576, 623, 645/Binding site: carbohydrate (Asi	
Query Match 94.0%; Score 15052; DB 1; Length 3010;	
Best Local Similarity 93.2%; Pred. NO. 0;	
Matches 2804; Conservative 87; Mismatches 119; Indels 0; Gaps 0;	
QY 1 MSTNPKPORKTKRNTNRRPQDVKPGGGQIVGGVYLLPRRGRPLGVRATRKASERSOPRG 60	
Db 1 MSTNPKPORKTKRNTNRRPQDVKPGGGQIVGGVYLLPRRGRPLGVRATRKTSERSQPRG 60	
QY 61 RRQPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGLLSPRGRSPSWGPTDPRRSRNILG 120	
Db 61 RRQPIPKARQPEGRAWAQPGYPWPLYGNEGLGWAGLLSPYGSRRPWGPTDPRRSRNILG 120	
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVEDGVNYATGNLPGCSFIFLIA 180	
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVEDGVNYATGNLPGCSFIFLIA 180	
QY 181 LLSCLTI PASAYEVNRVNSGIYHVNTDCNSSIIVTAAADVIHMTPCVCVQNGNSRCWV 240	
Db 181 LLSCLTVPSVTEYVRNVSIGVHVNTDCNSSIIVTAAADVIHMTPCVCVQNGNSRCWV 240	
QY 241 ALTPTLAARNASVPTTITRRHVDLIVGTAAAFCSAMVYVGLDLCGSIPLVSOLFTRSPRRHET 300	
Db 241 ALTPTLAARNASVPTTITRRHVDLIVGTAAAFCSAMVYVGLDLCGSIPLVSOLFTRSPRRYET 300	
QY 301 VQDCNCSYPGHVSGHRMAWDMMMNWSPTTALVWSQLLRIPQAVVDMVAGHWGLAGLA 360	
Db 301 VQDCNCSLYPGHVSGHRMAWDMMMNWSPTTALVWSQLLRIPQAVVDMVAGHWGLAGLA 360	
QY 361 YYSVMGNWAKVLIYVALLFAGVDGTHTTGRVAGHTTSFTSLFSSGASOKIOLVNTNGSW 420	
Db 361 YYSVMGNWAKVLIYVALLFAGVDGTHTTGRVAGHTTSFTSLFSSGASOKIOLVNTNGSW 420	
QY 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCPERMASCRPIDWFAQWGPITYTKPNS 480	
Db 421 HINRTALNCNDSINTGFFAALFYVKKFNSSGCSERMASCRPIDRFAQWGPITYHAESRS 480	
QY 481 DQRPYCMHYAPRPCGVVPASQVCGPVYCTFTSPVVGTTDRSGVPTYSGWNETDVMNLN 540	
Db 481 DQRPYCMHYAPQCGIVPALQVCGPVYCTFTSPVVGTTDRSGVPTYSGWNETDVMNLN 540	
QY 541 NTRPPQGNWFCTWNSTGFTKTCGGPRCNITGGVGNRTLI CPTDCFRKHPEATYTKCSG 600	
Db 541 NTRPPQGNWFCTWNSTGFTKTCGGPRCNITGGVGNRTLI CPTDCFRKHPEATYTKCSG 600	
QY 601 PMLTPRCVLDVYPRYLWHPYCTLNFTNFIFKVMYVGVGEHRLNAACNWTGRGCRNLEDSDRS 660	
Db 601 PMLTPRCVLDVYPRYLWHPYCTLNFTNFIFKVMYVGVGEHRLNAACNWTGRGCRNLEDSDRS 660	
QY 661 ELSPLLLSTTETWQILPCAFPTTLPALSTGLIHLHQNIVDQVLYGVGSAFVSPFAIKWEYIL 720	
Db 661 ELSPLLLSTTETWQILPCAFPTTLPALSTGLIHLHQNIVDQVLYGVGSAFVSPFAIKWEYVL 720	

QY 721 LFLLLADARVACACLLWMMLLIAQAEAALENLVINAASVAGAHGILSFLVFFCAAWYIKG 780
Db 721 LFLLLADARVACACLLWMMLLIAQAEAALENLVINAASVAGAHGILSFLVFFCAAWYIKG 780
QY 781 RLAPGAAYAFYGVWPLLLLLLLALPPRAYALDREMAASCGAVLVGLVFLTSPYKVELT 840
Db 781 RLVPGAAYAFYGVWPLLLLLLLALPPRAYALDREMAASCGAVLVGLVFLTSPYKAFILA 840
QY 841 RLILWMLQYFTTTRAEAHQWVWVPLNVVRGGRDAIILLTCAVHPELIIDIITKLMLAILGPLM 900
Db 841 RLILWMLQYFTTRVEACQLQWVPLNVVRGGRDAIILLTCAVHSELIIDIITKILLAILFGPLM 900
QY 901 VLOAGITRVEFYFVRAOGLIRACMLVRKVAGGHYVQVMFKLALGTGYVYNHULTPLRDWA 960
Db 901 VLOAGLTRVEFYFVRAOGLIRACMLVRKVAGGHYVQVMFKLALGTGYVYVYDHLTPLRDWA 960
QY 961 HAGLRDLAVAVEPVFSAMETKVIITWGAADTAACGDIILGLPVARSARGKEIFLGPADSLEG 1020
Db 961 HSGLRDLAVAVEPVFSAMETKVIITWGAADTAACGDIILGLPVARSARGKEILLGPADGPRE 1020
QY 1021 QGWELLAPITAYSOOTRGLVGLCIITSLTGRDKNOVEGEVQVSTATQSFATCINGVCWT 1080
Db 1021 QGWELLAPITAYSOOTRGLVGLCIITSLTGRDKNOVEGEVQVSTATQSFATCINGVCWT 1080
QY 1081 VYHAGSKTLTAGPKPITOMYTNVDLVLGVQWAPPGARSMTPCGSSGSDLYLVTRHADVI 1140
Db 1081 VYHAGSKTLTAGPKPINQMYTNVDQDLVGWQAPSGNASLTPCTYSSDLYLVTRHADVI 1140
QY 1141 PVRRRGDSRGLLSPREVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRRGDSRGLLSPREVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVES 1200
QY 1201 METTMRSPVTDNSTPPAVPQTFQVAHLHAPTGSCKSTKUPAAYAAQYKVLVLPNSVAA 1260
Db 1201 METTMRSPVTDNSTPPAVPQTFQVAHLHAPTGSCKSTKUPAAYAAQYKVLVLPNSVAA 1260
QY 1261 TLGFGAYMSKAHGIDPNIRGTVRTITGGSTYTYGKFLADGGCSGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGVDPNISITGVRTITGGSTYTYGKFLADGGCSGAYDIIICDECHS 1320
QY 1321 TDSTTILGIGTVLDOAETAGARLVLATATPPGSVTVPHNIEBIGLSNNGEIPFYGKAI 1380
Db 1321 TDSTSIIGIGTVLDOAETAGARLVLAATAPPGSVTVPHNIEBEVALPNTGEIPFYGKAI 1380
QY 1381 PIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVAYYRGLDVSUIPIGIDVVVVATDAL 1440
Db 1381 PLETIKGRHLIFCHSKKKCDLAALKSALGNNAVAYYRGLDVSUIPTSGDVVVVVATDAL 1440
QY 1441 MTGFTGDPDSVIDCNTCTQTVDFSLDPTFTIETTTVPQDAVSRSORGRGTGRSGIYR 1500
Db 1441 MTGVTGDPDSVIDCNTCTQTVDFSLDPTFTIETTTLPQDAVSRSORGRGTGRSGGIYR 1500
QY 1501 FVTGERSGFMFDSVLCEDYDAGAWYELTPAETSVRRLAYLNTPGLPVQCQDHLFEWES 1560
Db 1501 FVTGERSGFMFDSVLCEDYDAGAWYELTPAETSVRRLAYLNTPGLPVQCQVHLEWES 1560
QY 1561 VFTGLTHIDAHFLSOTKAGDNFPYLVAQATVCARAQAPPPSDQWKKLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSOTKAGENFPYLVAQATVCARAQAPPPSDQWKKLIRLKPTLHG 1620
QY 1621 PTPLLYRLGAVONEVILTHPTIKYIMACMSADLEVTWTWLVGGVLAALAAAYCLITGVS 1680
Db 1621 PTPLLYRLGAVONEVILTHPTIKFTIMACMSADLEVTWTWLVGGVLAALAAAYCLITGVS 1680
QY 1681 VIVGRIILSGRPVDPREVLYOEDEMEECASOLPYIEQGMQLABQFKQALGLLOTAT 1740
Db 1681 VIVGRIILSGRPVDPREVLYOEDEMEECASHLPYIEQGMQLABQFKQALGLLQTAS 1740
QY 1741 KOAEAAAAPVESKWALETFFWAKHWNFTISGQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KOAEAAAAPVESKWALEAFWAKHWNFTISGQYLAGLSTLPGNPAIVSLMAFTASITSP 1800

QY 1801 LTTQNTLLFNILGVMWAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDIILAGYAGVA 1860
Db 1801 LTTQNTLLFNILGVMWAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDIILAGYAGVA 1860
QY 1861 GALVAFKVMGSGVPSTEDLVNLLPAILSPGALVVGVCVAAIILRRHVGPGEVQWNNRLI 1920
Db 1861 GALVAFKVMGSGVPSTEDLVNLLPAILSPGALVVGVCVAAIILRRHVGPGEVQWNNRLI 1920
QY 1921 AFASRGNHVSPTHYVPESDAAAARVTQILSSLLITQILKRLHOWINEDCTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHYVPESDAAAARVTQILSSLLITQILKRLHOWINEDCTPCSGSWLRDV 1980
QY 1981 WDWICTVLTDFTKWLQSKLLPRLPGVPPFLSCORGKVGWVRGDIQIMOTTCPCQAQIAGHVK 2040
Db 1981 WDWICTVLTDFTKWLQSKLLPRLPGVPPFLSCORGKVGWVRGDIQIMOTTCPCQAQIAGHVK 2040
QY 2041 NGSMRIVGPRCTCNTWHGTFPINAYTTGCTPSPAPNYSRALMRVAAEEVVEVTRVGDFF 2100
Db 2041 NGSMRIVGPRCTCNTWHGTFPINAYTTGCTPSPAPNYSRALMRVAAEEVVEVTRVGDFF 2100
QY 2101 YVTGMTDNDVKPCQVPAPPEFFTEVDGVRHLHRYAPACKPILLREDVTFOVLNQYLVSQ 2160
Db 2101 YVTGMTDNDVKPCQVPAPPEFFTEVDGVRHLHRYAPACKPILLREDVTFOVLNQYPVGSQ 2160
QY 2161 PCPEPEDVTVLTSMLTDPHSITAEAKRLARGSPSLASSASQLSAPLSKATCTTHD 2220
Db 2161 PCPEPEDVTVLTSMLTDPHSITAEAKRLARGSPSLASSASQLSAPLSKATCTTHD 2220
QY 2221 SPADALIEANLLWRQWGGNITRVESENKVILDSFEPLHAEGDEREISVAAAILKRSK 2280
Db 2221 SPADALIEANLLWRQWGGNITRVESENKVILDSFEPLHAEGDEREISVAAAILKRSK 2280
QY 2281 FPSALPIWARPDYNNPILLESWKDPDYVPPVHVCPLPPTKAPPIPPRRKRTVVLTESNV 2340
Db 2281 FPSALPIWARPDYNNPILLESWKDPDYVPPVHVCPLPPTKAPPIPPRRKRTVVLTESNV 2340
QY 2341 SSALABELATKTFGSSGSSAVDSGTATAPDQPSDDGDRGSDDESYSMPPLLEGEPGPD 2400
Db 2341 SSALABELATKTFGSSGSSAVDSGTATAPDQPSDDGDRGSDDESYSMPPLLEGEPGPD 2400
QY 2401 SDGSWSTVSEAESEDVCCSMSTYWTGALITPCAAEESKLPINPLNSLRHHNMVYAT 2460
Db 2401 SDGSWSTVSEAESEDVCCSMSTYWTGALITPCAAEESKLPINPLNSLRHHNMVYAT 2460
QY 2461 SRASILROKQKVTDFRLQVLDDHYRDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
Db 2461 SRASILROKQKVTDFRLQVLDDHYRDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
QY 2521 GYKAKOVNRNLSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVPCVQPEKGRKPARLIV 2580
Db 2521 GYKAKOVNRNLSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVPCVQPEKGRKPARLIV 2580
QY 2581 FPDGLGVRCMKALYDVVWSTLPOAVMGSSYGFQYSKQRFVEFLVNTWKSCKCPMGFSYDT 2640
Db 2581 FPDGLGVRCMKALYDVVWSTLPOAVMGSSYGFQYSKQRFVEFLVNTWKSCKCPMGFSYDT 2640
QY 2641 RCPDSTVTESDIRVEESIYOCCLDAPARQAIRSLTERLYIGGPLTNSKQNCQYRCRA 2700
Db 2641 RCPDSTVTESDIRVEESIYOCCLDAPARQAIRSLTERLYIGGPLTNSKQNCQYRCRA 2700
QY 2701 SGVLTTSCGNLTLCYLKATACRAAKLQDCTMLVNGDDLVVVCESAGTQEDAAALAPTE 2760
Db 2701 SGVLTTSCGNLTLCYLKATACRAAKLQDCTMLVNGDDLVVVCESAGTQEDAAALAPTE 2760
QY 2761 AMTRYSAPPDPPQPEYDLELITSCSNVSVVAHDASGRVYLYTRDPTTPPLARAAMETAR 2820
Db 2761 AMTRYSAPPDPPQPEYDLELITSCSNVSVVAHDASGRVYLYTRDPTTPPLARAAMETAR 2820
QY 2821 HTPNWSWLGNIIMYAPTILWARMILMTHFFSILLAQOLEKALDQICQACYSIPLDUPQ 2880
Db 2821 HTPNWSWLGNIIMYAPTILWARMILMTHFFSILLAQOLEKALDQICQACYSIPLDUPQ 2880
QY 2881 IIERLHGLISAFTHLSYSGEINRVASCLRKLGVPLTRWHRARSVRAKLLSQGGAATC 2940


```
Db 2881 I|QRLHGLS|AFSLHSGPEINRVASCLRLKGLVPLRVHRPARRVRAKLLSQGGRAATC 2940
Qy 2941 GRYLFNFAVRTKLKLTPIPAASQOLDLSGWFAVAGYGGDIYHLSLRARPRWPLCLLLLSV 3000
Db 2941 GRYLFNFAVRTKLKLTPIPAASQOLDLSGWFAVAGYGGDIYHLSLRARPRWFWMLLLLSV 3000
Qy 3001 GVGIIYLLPNR 3010
Db 3001 GVGIIYLLPNR 3010

RESULT 6
GNWVC3
Name: polyprotein - hepatitis C virus (strain HCV-1)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain HCV-1)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: A39166; P00403; P00404
R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coi
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A:Title: Genetic organization and diversity of the hepatitis C virus.
A:Reference number: A39166; MUID:91172826; PMID:1848704
A:Accession: A39166
A:Molecule type: mRNA
A:Residues: 1-3011 <CHO>
A:Cross-references: UNIPROT:P26664; UNIPARC:UPI0000131E19; GB:M62321; NID:G329873; PIDN:
R.Chan, S.W.; McMahon, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I
J. Gen. Virol. 73, 1131-1141, 1992
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A:Reference number: P00393; MUID:92268871; PMID:1316939
A:Accession: P00403
A:Molecule type: genomic RNA
A:Residues: 1577-1633 <CHA>
A:Cross-references: UNIPARC:UPI0000174A00; DDBJ:D10128
A:Experimental source: isolates E-b16
A:Accession: P00404
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1577-1633 <CH2>
A:Cross-references: UNIPARC:UPI0000174A00
A:Experimental source: isolates E-b17
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match 88.2%; Score 14119.5; DB 1; Length 3011;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 2564; Conservative 226; Mismatches 220; Indels 1; Gaps 1;

Qy 1 MSTNPKPQKRTNRRNQDVKFGCGQIVGGVILLPRGRPLGVRATRKASERSQPRG 60
Db 1 MSTNPKPQKKNRNRNQDVKFGCGQIVGGVILLPRGRPLGVRATRKTSERSQPRG 60
Qy 61 RROQIPKARRPGRWAQPCYFWPLVYGNELGWAGWLLSPGRSRPSWGPTDPRRSRLG 120
Db 61 RROQIPKARRPGRWAQPCYFWPLVYGNELGWAGWLLSPGRSRPSWGPTDPRRSRLG 120
Qy 121 KVIDLTCTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFIFLLA 180
Db 121 KVIDLTCTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFIFLLA 180
```

```
Db 121 KVIDLTCTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFIFLLA 180
Qy 181 LLSCLTTPASAEYVRNVSGIYHVNTDCNSNSI VVEAADVI MHTPGCVPCVOGNSRCWV 240
Db 181 LLSCLTTPASAEYVRNVSGIYHVNTDCNSNSI VVEAADVI MHTPGCVPCVOGNSRCWV 240
Qy 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAPCSAMYVGDLCGSIPLVSQLTFSPRRHET 300
Db 241 AMTPVTATRDGKLPATQLRRHIDLLVGSATLCSALYVGDLCGSIPLVSQLTFSPRRHET 300
Qy 301 VDCNCSIYPGHVSGHRMAMDMNWSPTTALVVSOLLRIPOAVVDMVAGAHGWGLAGLA 360
Db 301 TQGCNCSIYPGHITGHRMAMDMNWSPTTALVVAQLRIPOAILDMTAGAHGWGLAGLA 360
Qy 361 YSVWGNWAKLIVALLFAGVDGTHTTGRVAGHTSGFTSLFSSGASQKQLVNTNSGW 420
Db 361 YFSWGNWAKLVLLVLLFAGVDAETHVTGSGAGHTVSGFVSELLAPAKQNVQLINTNSGW 420
Qy 421 HINRTALNCNDSLOTFFAALFYAHKFNSSGCPERMASCRPIDMFAQOGPIITYTKPNSS 480
Db 421 HLNSTALNCNDSLANTGLAGLFYHHKFNSSGCPERLASCRPLTDFDQCGWPISYANGSGP 480
Qy 481 DORPYCHYAPRPGCVVPASOVCPVYCFTPSPVVGTTDRSGVPTYSWGENETDMLLN 540
Db 481 DORPYCHYPPKPGCVIPAKSVCPVYCFTPSPVVGTTDRSGVPTYSWGENETDVFVLN 540
Qy 541 NTRPPQGNWFCCTWNSTGFTKTCGGPPCNIGVGNRTLCPTDCFRKHPEATYTKCGSG 600
Db 541 NTRPPLGNWFCCTWNSTGFTKVCAGPCVJGGAGNNTLHCTDCFRKHPEATYTKCGSG 600
Qy 601 PWTLPRLVDYPRYLWHYPCPLNFISFKRMVYGVGEHRLNAACNWTGERCNLEDRDRS 660
Db 601 PWITPRLVDYPRYLWHYPCPLNFISFKRMVYGVGEHRLNAACNWTGERCNLEDRDRS 660
Qy 661 ELSPLLLSTTQWQILPCAFITLPAISLGLIHLHQNIVDVQYLYGVGSFAVFAIKWEYIL 720
Db 661 ELSPLLLTTQWQVLPSCFTLPAISLGLIHLHQNIVDVQYLYGVGSFAVFAIKWEYIV 720
Qy 721 LFLFLLDARVCACIMWMLLIQAQAEALENAVLNAAVAGAHGLSLFLVFCAAWYIKG 780
Db 721 LFLFLLDARVCSCIMWMLLIQAQAEALENAVLNAAVAGAHGLSLFLVFCAAWYIKG 780
Qy 781 RLAPGAAAYFYGVWPLLLALLPRAVALDREMAASCGGAVLVGLVFLTLSPYKVFIL 840
Db 781 KWPFGAVVTFYGMWPLLLALLPRAVALDREMAASCGGAVLVGLVFLTLSPYKRVIS 840
Qy 841 RLIIWLVQVITRAAHQVWVPPNVRGGRDAIILLTCAVHPELIDITKULLAILGLPLM 900
Db 841 WCLMWLVQVITRAVQAHLVWVPPNVRGGRDAIILLTCAVHPELIDITKULLAILVFGPLW 900
Qy 901 VLOAGITRVYFVRAQGLIRACMLVRKVAGGHYQVWFMKLGALTGTYYVNHLTPLRDA 960
Db 901 ILQASLLKVPYFVRVQGLIRFALARKMIGHYQVWYIILKLGALTGTYYVNHLTPLRDA 960
Qy 961 HAGLRDLAVAPVVPVFSAMETKVIITWGAADTAACGDIILLGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HNGLRDLAVAPVVPVFSAMETKVIITWGAADTAACGDIINGLPVSARRGKEIFLGPADGVS 1020
Qy 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTRDKNVQVEGVVSTATQSFATCINGVCWT 1080
Db 1021 KGWRLLAPITAYSQOTRGVLGCIITSLTRDKNVQVEGVVSTATQSFATCINGVCWT 1080
Qy 1081 VYHAGSKTLAGPKGPIQWYTNVDLDLVGWAQPPGASMTPCSGSSDLVLTTHADVI 1140
Db 1081 VYHAGSKTLAGPKGPIQWYTNVDLDLVGWAQPPGASMTPCSGSSDLVLTTHADVI 1140
Qy 1141 PVRRGDSRGLSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAAVCTRGVAKAVDIPVES 1200
Db 1141 PVRRGDSRGLSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAAVCTRGVAKAVDIPVEN 1200
Qy 1201 METWRSVFTDNTSTPPAVPOTFOVAHLHAPTGGSKSTKVPAAVAAQYKVLVLPNSVAA 1260
Db 1201 LETWRSVFTDNTSTPPAVPOTFOVAHLHAPTGGSKSTKVPAAVAAQYKVLVLPNSVAA 1260
```

```
QY 1261 TLGFGAYMSKAHGDNPRTITGGITVSTYKFLADGCGSGAYDIIICBECHS 1320
Db 1261 TLGFGAYMSKAHGDNPRTITGGITVSTYKFLADGCGSGAYDIIICBECHS 1320
QY 1321 TDSTTILIGITVLDOAFTAGARLVLATATPGSVTPPHNIEIIGLSNNGEIPFYGKAI 1380
Db 1321 TDATSILIGITVLDOAFTAGARLVLATATPGSVTPPHNIEIIGLSNNGEIPFYGKAI 1380
QY 1381 PIBAIGKGRHLIFCHSKKKCDLAALKTLGLGNNAVAYRGDLSVPIPIGDVVVATDAL 1440
Db 1381 PLEVIGKGRHLIFCHSKKKCDLAALKVALGINAVAYRGDLSVPIPTSGDVVVATDAL 1440
QY 1441 MTGFTGDFSDVIDCNTCVTQTVDFSLDPTFTTETVTPQDAVRSQRRGRTGRSGIYR 1500
Db 1441 MTGYTGDFSDVIDCNTCVTQTVDFSLDPTFTTETITLTPQDAVSKTQRRGRTGRGKPGIYR 1500
QY 1501 FVTPGERPSGFMFSSVLCBVDAGCAWYELTPARTSVRLRAYLNTPLGVPQDHLFEWES 1560
Db 1501 FVAPGERPSGFMFSSVLCBVDAGCAWYELTPAETTVRLRAYMNTPLGVPQDHLFEWEG 1560
QY 1561 VFTGLTHIDAFHLSQTKQAGNFPYIVAYQATVCARAQAPPPSDQWKKLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAFHLSQTKQSGENLPYLVAQATVCARAQAPPPSDQWKKLIRLKPTLHG 1620
QY 1621 PTPLLYRLGAVONEVILTHPTTKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGVS 1680
Db 1621 PTPLLYRLGAVONEVILTHPTTKYIMTCMSADLEVTSTWLVGGVLAALAAAYCLSTGCV 1680
QY 1681 VIVGRIILSGKPAVVPDREVLVYQBFDEMEECASOLPYIEQGMQIAEQFKKALGLLQAT 1740
Db 1681 VIVGRVVLGKPAIIPDREVLVYREFDEMEECQHLPIYIEQGMMLAEQFKKALGLLQAT 1740
QY 1741 KOAEAAAPVVEKWALETFWAKHMNFIISGQYLAGLSTLPGNPATIASMAFTASTSP 1800
Db 1741 ROAEVIAPAVQTNQKLETFWAKHMNFIISGQYLAGLSTLPGNPATIASMAFTAAVTSP 1800
QY 1801 LTTQNTLNFNLGWWAAQLAPPSAASFVAGAGTAGAAGVIGIGLVLDILAGYGAGVA 1860
Db 1801 LTTSTQLTNFNLGWWAAQLAPGAATAFVAGAGLAGAAGVIGLVLDILAGYGAGVA 1860
QY 1861 GALVAFKMSGEVSTEDLVNLLPAILSPGALVGVWCAALRRHVGPGECAVOMNRLLI 1920
Db 1861 GALVAFKMSGEVSTEDLVNLLPAILSPGALVGVWCAALRRHVGPGECAVOMNRLLI 1920
QY 1921 AFASRGNHVSPTHVVPESDAAARVTOILSSLTITQLKRLHQWINEDCSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHVVPESDAAARVTAIILSSLTITQLKRLHQWISSECTTFCSGSWLRDI 1980
QY 1981 WDWICTVLTDFTKWLQSKLLPRLPGVPLFSCQRYKGYWRGDGIMQTTCPCGAIAGHVK 2040
Db 1981 WDWICEVLSDFTKWLKAKLMPQLPGIPFVSCQRYKGYWRVDGIMHTRCHGAEITGHVK 2040
QY 2041 NGSMRIVGPRCSNTHGTFPINAYTTGCTPSPAPNYSRALWRVAEYVEVTRVGDGFH 2100
Db 2041 NGTMRIVGPRCSNTHGTFPINAYTTGCTPLPAPNYTFALWRVSAEYVEIRQVGDGFH 2100
QY 2101 VYTGMTTNDVKPCQVPAPEFTEVDGVLRLHRYAPACKPLAREDTVFQVGLNQVLGSOL 2160
Db 2101 VYTGMTTNDLKCPCQVSPPEFTELDGVLRLHRAFPPCKPLREEVSRVGLHEYPVGSOL 2160
QY 2161 PCEPEPDVTLTSMLTDPESHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTHHD 2220
Db 2161 PCEPEPDVAVLTSMLTDPESHITAEAGRLARGSPPSVASSASQLSAPSLKATCTAHD 2220
QY 2221 SPDADLIEANLLWRQEMGNITRVESENKVVILDSFEPHLHAEGBDEREISVAEILKRKR 2280
Db 2221 SPDAELIEANLLWRQEMGNITRVESENKVVILDSFDPPLVAEEDEREISVPAEILKRKR 2280
QY 2281 FPSALPIWARPDYNPPLLESKWDPPYVPPVHGCPLPPTKAPPIPPPRKRKTIVLTESNV 2340
Db 2281 FAQALPVWARPDYNPPLVETMKKPDYEPVPPVHGCPLPPPKSPVPPPRKRKTIVLTSTL 2340
```

```
QY 2341 SSALAEATKTTGSSGSSAVDSCTATAPDLASDDGDKGSDVESYSMPPLEGEPDPL 2400
Db 2341 STALAEATRSFGSSSTSGITGDNITTSSEBAPSGCPDSDAESYSMPPLEGEPDPL 2400
QY 2401 SDGSMWTSSEEA-SEDDVCCSMSTWTGALITPCAEESEKLPINPLSNLSLRHHNMVYAT 2459
Db 2401 SDGSMWTSSEANAEDVCCSMSTWTGALVTPCAAEEOKLPINALSNLSLRHHNLVYST 2460
QY 2460 TSSASLRQKKVTFDRLQVLDLHDHYRDLVKEMKAKASTVAKLLSIEEACKLTPPHSAKSK 2519
Db 2461 TSSACORQKKVTFDRLQVLDLSDHYQDLKEVKAASKVKANLLSVEEACSLTPPHSAKSK 2520
QY 2520 FGYGAKDVRLNSRAVNHRSVWEDLEDETETPIDTITMAKSEVFCVQPSKGRKPARLI 2579
Db 2521 FGYGAKDVRCARKAVTHINSVWMDLLEDNVTPIDTITMAKNEVFCVQPSKGRKPARLI 2580
QY 2580 VFPDLGVRCCKMALVDVSTLPOAVMGSSYGFQYSPKORVEFLVNTWTKSKKCPMGFSYD 2639
Db 2581 VFPDLGVRCCKMALVDVTKLPLAVMGSSYGFQYSPGQRFVFLVQAWKSKKTPMGFSYD 2640
QY 2640 TRCFDSTVTESDIRVBESYIQCCLDAPAEARQAIRSLTERLYIGG?LTNSKGQNGYRRCR 2699
Db 2641 TRCFDSTVTESDIRTEEAIIQCCLDLPQARVAIKSLTERLYVGGPLNSRGNGCYRRCR 2700
QY 2700 ASGLVTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVTICESAGTOEDAAALRAFT 2759
Db 2701 ASGLVTTSCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLVTICESAGVQEDAAALRAFT 2760
QY 2760 EAMTRISAPPDPPQPEYDLELITSCSNVSVVAHDASGKRVYVLTDRPTTPLARAAWETA 2819
Db 2761 EAMTRISAPPDPPQPEYDLELITSCSNVSVVAHDGAGKRVYVLTDRPTTPLARAAWETA 2820
QY 2820 RHTPINSWLGNIIMYAPTLWARMLTMHFFSILLAOBLEKALDCCIYGCYSTIEPLDLP 2879
Db 2821 RHTPINSWLGNIIMFAPTLWARMLTMHFFSVLIARQQLDCEIYGCYSTIEPLDLP 2880
QY 2880 QIITERLHGLSAFTLHYSVSGEINRVASCLRGVPLRTWRHARSVRAKLLSGGGAAT 2939
Db 2881 PIQRLHGLSAFSLHYSVSGEINRVACLRKLGVPPLRANWRHARSVRARLLARGGAAI 2940
QY 2940 CGRYLFENAVRTKLKLTPIPAASQLDLSGMFVAGYSGGDIYHLSRSRPRWFPCLLILLS 2999
Db 2941 CGKYLFNNAVRTKLKLTPIPAAGQLDLSGMFTAGYSGGDIYHVSHPRPMWFCLLLA 3000
QY 3000 VGVGIYLLPNR 3010
Db 3001 AGVGIYLLPNR 3011
```

RESULT 7

S40770

genome polyprotein - hepatitis C virus

N:Contains: capsid protein C; envelope protein M; hepatitis virus (EC 3.4.21.98) (nonstructu

protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004

C:Accession: S40770; PID:285

R:Okamoto, H.

submitted to the EMBL Data Library, March 1992

A:Reference number: S40770

A:Accession: S40770

A:Molecule type: genomic RNA

A:Residues: 1-3011 <OKA>

R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda,

Jpn. J. Exp. Med. 60, 167-177, 1990

A>Title: The 5'-terminal sequence of the hepatitis C virus genome.

A:Reference number: PC1284; MUID:91013116; PMID:2170712

A:Accession: PC1285

A:Molecule type: genomic RNA

A:Residues: 1-513 <OK2>

A:Cross-references: UNIPARC:UPI00000F4078; GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:

A:Experimental source: isolate HC-J1

C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serine
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus nonstructural protein NS3 #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2011/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3013/Product: nonstructural protein NS5 #status predicted <NS5>

```
QY 1921 AFASRGNHVSPTHYVPESDAAARVTOILSSLTITQLKELHOWINBDCSTPCSGSWLRDV 1980
DB 1921 AFASRGNHVSPTHYVPESDAAARVTAISSLTVTQLRLHQLWSSESTPCSGSWLRDI 1980
QY 1981 WDHICTVLTDFKTLWLOSKLLPRLPGVPFFLSQBGYKGVWRGIGIMOTTCCPQAGQIAGHYK 2040
DB 1981 WDWICEVLSDFKTLWTKLMPHLPFGIPFVSCQBGYKGVWRGIGIMTRCHCGAEITVGHYK 2040
QY 2041 NGSMRVGPRCTCNTHGTFPIINAYTTGCTCPAPNYSRALWRVAEEVVEVTRVGDHF 2100
DB 2041 NGTMRIVGPKTRMWSGTFPIINAYTTGCTCPAPNYTFALWRVSAEEVVEIRRVGDHF 2100
QY 2101 YVTGMTTNNVKCPQVPAPFEFTFVGVRLHRYAPACKPLLRBDVTFQVGLNOYLVGSQL 2160
DB 2101 YVTGMTTNNVKCPQVPAPFEFTFVGVRLHRYAPACKPLLRBDVTFQVGLNOYLVGSQL 2160
QY 2161 PCPEPDDVTLTSMLTDPSSHITAEAKRLRLARGSPPLSSASSASOLSAPSLKATCTTHD 2220
DB 2161 PCPEPDDVTLTSMLTDPSSHITAAAGRLRLARGSPPEASASSASOLSAPSLKATCTINH 2220
QY 2221 SPDAELLEANLRQEMGNNITRVESENKVVILDSPEPLHAEGDEREISVAABILKRSK 2280
DB 2221 SPDAELLEANLRQEMGNNITRVESENKVVILDSPEPLVAEEDEREISVPAELKRSR 2280
QY 2281 FPSALPIWARPDVNPPLLESWKPDYVPPVHVHGCPLPPTKAPITPPPRKRTVVLTESNV 2340
DB 2281 FTQALPIWARPDVNPPLIETWKXNPVEPPVHVHGCPLPQPSPVPPPRKRTVVLTESTL 2340
QY 2341 SSALAEIATKTFGSSGSAVDSGTATLPLDASDDGKGDVRSYSNMPLEGPDPDL 2400
DB 2341 STALAEIATKTFGSSGSAVDSGTATLPLDASDDGKGDVRSYSNMPLEGPDPDL 2400
QY 2401 SDGSWSVSEEA-SEDEVVCCSMVWTGALITPCAAEESKLPIPNPLNSLLRHNNMYAT 2459
DB 2401 SDGSWSVSEAGTEDVVCMSMYTWTGALITPCAAEEOQLPINALSLLRHNNLVYST 2460
QY 2460 TSRSASLRQKVTFRDLQVLDDHVRDVLKEMKAASTVKAKLISIEACKLTPPHSAKS 2519
DB 2461 TSRSACORQKVTFRDLQVLDDHVRDVLKEMKAASTVKAKLISIEACKLTPPHSAKS 2520
QY 2520 FGYGAKDVRLNLSRAVNHRSWEDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLI 2579
DB 2521 FGYGAKDVRCARAVNHNSVWKLDESDVTPITQTTIMAKNEVFCVQPEKGRKPARLI 2580
QY 2580 VFPDLGVRVCEKMAIYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNTWKSCKPWSYD 2639
DB 2581 VFPDLGVRVCEKMAIYDVVSTLPQAVMGSSYGFQYSPQORVEFLVQAWKSKRTPMGFSYD 2640
QY 2640 TRCFDSTVTESDIRVEESIYOCCLAPEAROAIIRSLTERLYIGGLPLTNSKGONCYRRCR 2699
DB 2641 TRCFDSTVTESDIRTEEAIIYOCCLDPOARVAIRSLTERLYIGGLPLTNSRGENCYRRCR 2700
QY 2700 ASGVLTTSCGNTLTCLYKATAACRAAKLQDCTMLVNGDDLVIICESAGTQDEAALRAFT 2759
DB 2701 ASGVLTTSCGNTLTCLYKATAACRAAKLQDCTMLVNGDDLVIICESAGVQDEAASLRAFT 2760
QY 2760 EAMTRYAPPDGPPOPEYDLELITSCSNVSVVAHDASGRVYVLTTRDPTPLAARAWETA 2819
DB 2761 EAMTRYAPPDGPPOPEYDLELITSCSNVSVVAHDGTCRKYVYLTTRDPTPLAARAWETA 2820
QY 2820 RHTPINSWLGNIIMYAPTILWARMILMTHFFSILLAQEOLEKALDQIYGCACYSIEPLDLP 2879
DB 2821 RHTPINSWLGNIIMYAPTILWARMILMTHFFSILLAQEOLEKALDQIYGCACYSIEPLDLP 2880
QY 2880 QTIERLHGLSAFTLHSGPEINRVASCLRKLGVPLRTWHRARSVRAKLLSQCGRAAT 2939
DB 2881 PIIOQLHGLSAFTLHSGPEINRVAAACLRKLGVPPLRAWRHARSVRARLLSRGGRRAI 2940
QY 2940 CGRYLFNNAVRKTLKLTIPPAASQOLDLSGFWAGYSGGDIYHSLSRARPRWFFPLCLLLLS 2999
DB 2941 CGRYLFNNAVRKTLKLTIPPAAGRLDLSGFWFTAGYSGGDIYHVSVAHRPRWFFWFLCLLLA 3000
```

```
QY 3000 VQVGIVLLPNR 3010
DB 3001 AGVGIVLLPNR 3011
RESULT 8
GNWVCH
genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructur
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36814; A41546
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus: c
A;Reference number: A36814
A;Accession: A36814
A;Molecule type: genomic RNA
A;Residues: 1-3011 <INC>
A;Cross-references: UNIPROT:P27958; UNIPARC:UPI000004ACD7; GB:M67463; NID:9329737; PIDN:I
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A;Title: Genomic structure of the human prototype strain H of hepatitis C virus: comparis
A;Reference number: A41546; MUID:92052256; PMID:1658800
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F;A-115/Product: capsid protein C #status predicted <CPC>
F;I16-191/Product: envelope protein M #status predicted <EPM>
F;I92-389/Product: major envelope protein E #status predicted <NEE>
F;I390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;I730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;I1007-1635/Product: hepacivirin #status predicted <NS3>
F;I230-1237/Product: nucleotide-binding motif A (P-loop)
F;I230-1237/Region: nucleotide-binding motif B
F;I312-1317/Region: nucleotide-binding motif B
F;I316-1319/Region: DEXH motif
F;I1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>
F;I1863-2013/Product: nonstructural protein NS5 #status predicted <N4B>
F;I2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;I196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23
Query Match 87.0%; Score 13935.5; DB 1; Length 3011;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2545; Conservative 226; Mismatches 239; Indels 1; Gaps 1;
QY 1 MSTNPKPQRTKNTNRRPODVKPPGGQIVGGQIIVGGYLLPRRGPRLGVRAIRKATKASERSQPRG 60
DB 1 MSTNPKPQRTKNTNRRPODVKPPGGQIVGGYLLPRRGPRLGVRAIRKATKASERSQPRG 60
QY 61 RRQPIKARPEGEAQAQPCYPMPLVYNEGLGWAGLLSPRGRSPSGPTDPRRRSRNLG 120
DB 61 RRQPIKARPEGEAQAQPCYPMPLVYNEGLGWAGLLSPRGRSPSGPTDPRRRSRNLG 120
QY 121 KVIDLTCGFADLMGYIPLVGAAPLGGAAARALAHGVRVLEGGVNYATGNLPCGSFSIFLLA 180
DB 121 KVIDLTCGFADLMGYIPLVGAAPLGGAAARALAHGVRVLEGGVNYATGNLPCGSFSIFLLA 180
QY 181 LLSCLTIPASAYEVRNVSGIYHVNTDCNSISVVEAADVIMHTPGCVPCVOEGNSSRCWV 240
DB 181 LLSCLTIPASAYVRNSSLGYHVNTDCPNSSVVEAADAILHTPGCVPCVREGNASRCWV 240
QY 241 ALTPFLAARNASVPTTIRRHVDLLVGTAAFCSSMYVGDLCGSIFLYSQIFTFSPRHET 300
DB 241 AVTPTVATRDGKLTPTQLRRHIDLLVGSATLCSALYVGDLCGSIFLYSQIFTFSPRHET 300
QY 301 VQDCNCSITYPGHVSGHAWMMNMWSPPTALVVSQLLRIPOAVVDMVAGAHGWAGLAGLA 360
DB 301 TQDCNCSITYPGHITGHRMAMNMNMWSPPTAALVVAQLLRIPOAIDMIDTAGHGWAGLAGIK 360
QY 361 YYSVMGVNNAKVLIVALLFAGVDGTHTTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW 420
```


QY	2580	VFPDLGVRVCEKALYDVVSTLPQAVMGSSYGFQYSPKQORVEFLVNTWKSKKCPMGFSYD	2639
Db	2581	VFPDLGVRVCEKALYDVVSKPLAVMGSSYGFQYSPGQORVEFLVQAKSKKTFMGLSYD	2640
QY	2640	TRCFDSTVTESDIRVESIYQCCDLAPAPARQAIRSLTERLYIGGPLTNSGQNGCYRRCR	2699
Db	2641	TRCFDSTVTESDIRTEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCYRRCR	2700
QY	2700	ASGLVLTSCGNTLTCVLTAKTAAACRAAKLQDCTMLVNGDDLVTICESAGTQEDAAALRAFT	2759
Db	2701	ASRLVLTSCGNTLRIKARAAACRAAGLQDCTMLVCGDDLVTICESAGVQEDDAASLRAFT	2760
QY	2760	EAMTRYSAAPPDPQPEYDELELTSCSSNVSAHDASGRKVVYLTTRDPTTPLARAAWETA	2819
Db	2761	EAMTRYSAAPPDPQPEYDELELTSCSSNVSAHDAGKRVYLTTRDPTTPLARAAWETA	2820
QY	2820	RHTPINSWLGNIIMYAPTLMWMLMTHFFSIIAQEQLEKALDQCIYGACYSIEPLDLP	2879
Db	2821	RHTPVNSWLGNIIMFAPTLWMLMTHFFSVLIARDQEQALNCEIYGACYSIEPLDLP	2880
QY	2880	QIIRLHGLSAFTLHSPGEINRVASCLKLGVPPLRTWRHRARSVRAXLLSOGGGAAT	2939
Db	2881	PIIORLHGLSAFTLHSPGEINRVAACTLKLGPPLRWRHRAWSVRARLLARGGKAAI	2940
QY	2940	CGRYLFNWAIVTKLTIPIPAASQDLDSGHFVAGYSGGDIYHLSRARPWFPLCLLLLS	2999
Db	2941	CGRYLFNWAIVTKLTIPIAAGRLDLSGWFTAGYSGGDIYHVSHPARPFWFCLLLLA	3000
QY	3000	GVGVIYLLPNR 3010	
Db	3001	AGVGIYLLPNR 3011	
RESULT 9			
JC5620			
Genome polyprotein - hepatitis C virus (isolate EUH1480)			
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (EC 3.4.21.98) (nonstructural protein NS4a); nonstructural protein NS4b; nonstructural protein NS5			
C;Species: hepatitis C virus			
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004			
R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.			
Biochem. Biophys. Res. Commun. 236, 44-49, 1997			
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant			
A;Reference number: JC5620; MUID:97366593; PMID:9223423			
A;Accession: JC5620			
A;Molecule type: mRNA			
A;Residues: 1-3014 <CHA>			
A;Cross-references: UNIPROT:O39928; UNIPARC:UPI0000174A01; GB:Y13184			
A;Experimental source: genotype 5a, which predominates in South Africa			
A;Note: the translation of the nucleotide sequence is not complete in this paper			
C;Superfamily: hepatitis C virus genome polyprotein			
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin			
F;2-115/Product: capsid protein C #status predicted <CPC>			
F;116-191/Product: envelope protein M #status predicted <BPM>			
F;192-389/Product: major envelope protein E #status predicted <MEE>			
F;384-408/Region: hypervariable #status predicted			
F;390-730/Product: nonstructural protein NS1 #status predicted <NS1>			
F;731-1007/Product: nonstructural protein NS2 #status predicted <NS2>			
F;1008-1616/Product: hepatitis C virus #status predicted <NS3>			
F;1231-1238/Region: nucleotide-binding motif A (P-loop)			
F;1313-1318/Region: nucleotide-binding motif B			
F;1317-1320/Region: DEXH motif			
F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>			
F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>			
F;2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>			
F;2210-2249/Region: interferon sensitivity determining #status predicted			
Query Match 78.1%; Score 12496; DB 1; Length 3014;			
Best Local Similarity 75.0%; Pred. No. 0;			
Matches 2268; Conservative 309; Mismatches 423; Indels 24; Gaps 6;			

QY	1	MSNPKPQBKTNTNRRPQDVKPPGGQIVGGVYLLPRGPRLGVRATRKASERSQPRG	60
Db	1	MSNPKPQBKTNTNRRPQDVKPPGGQIVGGVYLLPRGPRLGVRATRKASERSQPRG	60
QY	61	RRQPIPKARRPRGRWAQPGYPWPPLYGNEGLWAGWLLSPRGRSP?SWGPTDPRRRRNILG	120
Db	61	RRQPIPKARRPTGRSGWQPGYPWPPLYANBGLWAGWLLSPRGRPNWGPNDPRKSPNIG	120
QY	121	KVIDLTLCQFADLMGYIPLVGAPLGGARALAHGVRLVEDGVNYYATGNLPCCSFISFLA	180
Db	121	RVHTLTLCQFPLMGYIPLVGPGVGVSRALAHGVRLVEDGNGYATGNLPCCSFISFLA	180
QY	181	LLSCLTIPASAEVRNVSGIYHVNTDCSNSSIVYEADVIMHTPGCVPCVQEGNSRCWV	240
Db	181	LLMCLTVPASAFYRNASGVYHVNTDCPNSSIVYEADNLLHAPGCPVCLEDNVRCWV	240
QY	241	ALTPTLAAENASVPTTITIRRHVDLVGTAAFCSAMVVDLCCGSIFFVSOLFTEPSRRHET	300
Db	241	QIITPTLSAFSGAVTALLRRADVILAGGAFCALVVGACGALSIVGQWFTYKPRQHTT	300
QY	301	VQDCNCSIYPGHVSGHRMAWMMWMSPTTALVVSQLLRIPOAVVDMVAGAHGVLAGLA	360
Db	301	VQDCNCSIYSGHITGRMAWMMWMSPTTALLMAQLLRIPQVVIDIAGHGVLLAAA	360
QY	361	YYSWVGWNAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW	420
Db	361	YFASTANWAKVILVLFAGVDGRTHTVGTVGQGLKSLTSPFNPQORQLQFVNTNGSW	420
QY	421	HINRTALNCNDSIQTOFFFAALFYAHKFNSSGCPERMASCRPIDWFAQGWPYIY-TKPNS	479
Db	421	HINSTALNCNDSIQTOGFIAGLMAHKFNSSGCPERMSSCRPLAADFQDGWGTISYATTSGP	480
QY	480	SDORPYCWHYAPPCGVVPASQVCGVYCTPSPVTVGTTRDGRSGVPYSGENETDVMLL	539
Db	481	SDDKPYCWHYPPRCGVVPARDVCGVYCTPSPVTVGTTRDGRGCPYNNGSNETDILL	540
QY	540	NNTRPPQGNWFGCTWMNSTGTCTKCGPPCNCIGVGNRTLICPTDCFRKHPETATYKCGS	599
Db	541	NNTRPPAGNWFGCTWMNSTGTVKNCGAPCNCNLGPTGNNSLKCTDCFRKHPDATYKCGS	600
QY	600	GPWLTPRCLVDYPRYLWHYPTCLNFSIFKVMYVGGVEHRLNAAACNWTGERCNLEDRDR	659
Db	601	GPWLTPRCLVHYPRYLWHYCTVNYTIFKVMFIFGLEHLEAAACNWTGERCNLEDRDR	660
QY	660	SEISPLLLSTTEWOILPCATPTLLPALSTGLIHQNIQVQYLYGVGSFAFVSPFAIKWEYI	719
Db	661	AEISPLLLHTTTQWAILPCSFPTPALSTGLIHQNIQVQYLYGLSSISVMAVAKWEYI	720
QY	720	LLFLLLADARVCACLMMMLLIAQAEAALENLVINAASVAGAHGILSFVFFCAAMVYK	779
Db	721	MLVFLLLADARICTCLLLILLICQAEATCKNVIINAAAAAGNHGFFWGLLVVCLAMHVK	780
QY	780	GRLLAPGAAYAFYGVWPLLLALLLALPPRAYALDREMAASCAGVAVGLVFLTSLSPYKVL	839
Db	781	GRLVPGATYICLVGVWPLLLVRLRLRPHRALDSDSGGTVCGLVLIVITIFLTITGYKKKV	840
QY	840	TRLIWLQYFETTRAEAHQWVWVPLNVNRRGRDAIILLTCAVHPELIIDITKLLAILGLPL	899
Db	841	VLVNMWLOQYFIARVEALIHVWVPLQVKGGRDAVIMLTCLFHPALGFEITKILFGILGPL	900
QY	900	MVLQAGITRVPYFVRAQGLIRACMLVRKVAGGHYVQVFMKLGALTGTYYVNHHTPLRDW	959
Db	901	YLLQHSLSKVPYFLRARALLRLCLLAKHLVYGVQVQAALLHGLRTGTIYDHLAPMKDW	960
QY	960	AHAGLRDLAVAVEPVVPSMETKVIWGTADTAACGDIILGLPVSARRKEIFLGPADSLE	1019
Db	961	AASGLRELTVATEDIVFSAMETKVIWGTADTAACGNIILAVLPVSARRGREIFLGPADDIK	1020
QY	1020	QGWRLLLAPITAYSQQTRGVLCIITSLTGDRKKNQVEGVQVSTATQSFATCINGVCW	1079
Db	1021	TSWRLLLAPITAYAAQTRGVLCIIVLSLTGRDKNEAGEVQFLSTATQTFIGICINGVMW	1080
QY	1080	TVYHAGSKTLAGPKGPITQWYTNVDLVLVQWAPPGARSMTPCSCGSSDLVLTVRHADV	1139

1081	Db	TLFHAGSKTLAGKGVQVQYTNVDKDLVGMFSPGKGLTRCTCGSADLYLVTSHADV	1140
1140	Qy	IPVRRDGRSGLLSPRVSYLKSGSGCPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVE	1199
1141	Db	IPARRDGTTRASLLSPRPISYLKSGSGGPIMCTPSGHVGVFRAAVCTRGVAKALEFVPVE	1200
1200	Qy	SMETMRSVPFTDNSTPPAVPQTQFVAHLHAPTSGSGSKTKVPAAAYAAQGYKVLVLPNSVA	1259
1201	Db	NLETTMRSVPFTDNSTPPAVPHEFQVGHLLHAPTSGSGSKTKVPAAAYAAQGYKVLVLPNSVA	1260
1260	Qy	ATLGFQAYMSKAHGIDNIRITGVRTITTTGSGITYSTYTKFLADGGSGGAYDIIIODECH	1319
1261	Db	ATFGFQAYMSRAYGVDNIRITGVRTITTTGAGITVSTYTKGFADGGSGGAYDIIIODECH	1320
1320	Qy	STDSTTILGTGLVDQAEAGARLVILATATPGSIVTPHPNIEETGLSNNGEIPFVGKA	1379
1321	Db	SQDATTILGTGLVDQAEAGARLVILATATPGSVTTTPHPNIEEVALPSEGEIPFVGRA	1380
1380	Qy	IPIBAIKGGRHLIFCHSKKKDELAALKTLGLGNNAVAYRGLDVSIVPTIGDVVVVATDA	1439
1381	Db	LPVLIIKGRHLIFANQKKAKETAKQWNKPGKAVAYRGLDVAVIPATGDDVVVCTDA	1440
1440	Qy	LMTGFTGDFSVIDCNTCVTQTQVDFSLDPTFTTETTTVPQDAVSRSORRGTGRGRGIY	1499
1441	Db	LMTGFTGDFSVIDCNSAVTQTQVDFSLDPTFTTETTTVPQDAVSRSORRGTGRGRGIY	1500
1500	Qy	RFVTPGRPSGMDSSVLCBCYDAGCAWYELTPAETSVRLRAYLNTPGLPVCODHLEFWE	1559
1501	Db	RYVSSGSRPSGIFDVSVLCBCYDAGCAWYDLTPAETTVRLRAYLNTPGLPVCQEHLEFWE	1560
1560	Qy	SVFTGLTHIDAHFLSOTKQAGDNPPYLVAQVATCARAQAPPPSWDOMKCLRLKPTLH	1619
1561	Db	GVFTGLTNIDAHMLSQAKQGENPPYLVAQVATCVRAKAPPPSWDMMKMLCLPILT	1620
1620	Qy	GPTPELLYRLGAVQNEVILTHPIKTYIMACHSADLEVVVTTWLVGGVLAALAAAYCLTTGS	1679
1621	Db	GPTPELLYRLGAVQNEILTHPIKTYIMACHSADLEVVTTWLVGGVVAALAAAYCLTVGS	1680
1680	Qy	VVIUGRIILSKPAVDPREVLYQEPDEMEECASQLPIEQGHQLABOFKQKALGILLOTA	1739
1681	Db	VAIUGRIILSGRPAITPDREVLVYQOFDEMEECASLPYDVDEARAIAGQFKKVLGLIGTA	1740
1740	Qy	TKQEAAPVVESSKWRALETFWAKHMWNFTSGIOYLAGLSTPLGNPAIASLMAPTASITS	1799
1741	Db	GQKAEITLKPAATSNWSKAEQWAKHMWNFVSGIOYLAGLSTPLGNPAVATLMSFTAATS	1800
1800	Qy	PLTTQNTLLFNILGGWVAQAAPPASAASAFVAGIAGAAGVSGIGLKVLVDIILAGYGAGV	1859
1801	Db	PLTTHTQTLTNILGGWVASQIAPPTAATAFVVSWMAGAAVGNIGLGRVLIDILAGYGTGV	1860
1860	Qy	AGALVAFKWSGEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVCPGEGAVQWNRL	1919
1861	Db	AGALVAFKMGGERPTABELVNLNLPSTLCFPGALVGVVCAAVLRHRTGPEGAVQWNRL	1920
1920	Qy	IAPASRGNHVSPTHVYVESDAAAARVTOILASLTITQLLKRHOWINEDCSTPCSGWLRD	1979
1921	Db	IAPASRGNHGSPTHVYFETDSAKVTQLLSLTVLSLLKRLHTWIGEDYSTPCDGTWLR	1980
1980	Qy	VWDMICTVLTDPKTWLQSKLLPRLPGVPFPLSCQRYGKGVWRGDMGIMTTCPCGAQIAGHV	2039
1981	Db	IWDVWCTALTDKFKAWLQAKLLPOLPGVPFFPSQCKGYKGVWRGDMVNSTKCPCCGATISGHV	2040
2040	Qy	KNGSMRIVGPRCTCNTHWGTTPPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDF	2099
2041	Db	KNGTMRIVGPKLCSNTWQGTFFPINATTTPGSPVAPAPNYPKFAWRVGAADYAEVRRVGDY	2100
2100	Qy	HYVTGMTDNVCKPCQVPABEFFTEVDGRLHRVAPACKPLLRREDVTFQGLNQYLVGSO	2159
2101	Db	HYITGVTDQNLKPCQVPSEBFFTELDGVRILHRFAPPNCNPLLREEVTFVGLSHSVVVGSO	2160
2160	Qy	LPCEPEPDVTVLTSMLTDPSHITAEAKRRLARGSPPSLASSASQASLPASLKATCTTTH	2219

Db	2161	LPCEPDPDVTVLTSMLSDPAHITAETAKRLNRKNGSPPSLANSSASQLSAPSLKATCTIQG	2222
Qy	2220	DSPADLIEANLLWRQEMGNIITRVESENKVTILDSFEPHLHABGDREISVAAEILRKSR	2279
Db	2221	HHFPDADLLKANLLWRQMGNIITRVEAENKVEILDCFKPLKEEDDRREISVADCFKGP	2280
Qy	2280	KFPSALPIIWARPDYNPPLLESWKDPDYPVPVHGCPLPPTKAPPIPPRRKR-TVVLTS	2338
Db	2281	AFPPALPVWARPGDPLLETWKRPDYPQVWGVCPIPPAGPPVPPLPRKRKPMELSDS	2340
Qy	2339	NVSALAEIATKTF-----GSGSSAVDSGTATALPDLASDDGDKSDSVESYS	2386
Db	2341	TVSQVWADLADARFKVDTPSIEGQDSALGTTSSQHDG-----PEEKRD-----NSDAASYS	2392
Qy	2387	SMEPLGEPCDPLDSGSMSTVSEASESDVVCSCMSYTTWTGALITPCAAEESKLPINPLS	2446
Db	2393	SMEPLGEPCDPLDSGSMSTVSGE--DNVVCSCMSYTTWTGALITPCSAAEESKLPINPLS	2450
Qy	2447	NSLLRHNNMYYAITSRSASLRQKKVTFDRQLQVLDHHRDVLKEMKAKASTVKAKLJSIEE	2506
Db	2451	NTLLRHNLVYSTSSRSASGLRQKKVTFDRQLQVLDHHRVVDSEMKRLASKVKARLLPLEE	2510
Qy	2507	ACKLTTPHSAKSPFGYGAQDVRLSSRAVNHHSRWEDLLEDTETPIDTITIMAKSEVFCV	2566
Db	2511	ACGLTTPHSAKSPFGYGAKEVRLSKKALKHIEGVWQDLDLSDSDPTLPTTITIMAKNEVFV	2570
Qy	2567	QPEKGRKPARLLVFPDLGVRCERKALVDVWSTLPOAVMGSSYGFQYSPKQVREFLVNT	2626
Db	2571	EPKGGKPARLLVFPDLGVRCERKALVDVAOKLFTALMGPSYGFQYSPAQRVDFLLKA	2630
Qy	2627	WKSCKCPMGFSYDTRCFDSTVTEESDIYQCCDLAPEARQAIRSLTERLIYIGPLT	2686
Db	2631	WKSCKIPMAFSYDTRCFDSTITEHDIMTEESIIYQSCDLQPEARVAIRSLTQRLYCGGPMY	2690
Qy	2687	NSGQNGCYRRCRASGVLTTSCTNTLTCYLKATAACRAAKLQDCTMLVNGDDLWVTCESA	2746
Db	2691	NSGQQCGYRRCRASGVFTTSMGNTWCTYIKALASCRAAKLRDCTLLVCGDDLVAICESQ	2750
Qy	2747	GTOEDAAALRAFTEAMTRYSAPDGPQPEYDLELITSCSSNNVSVAHDASGKRVYVLT	2806
Db	2751	GTHEDEASARAFTEAMTRYSAPDGPVPVPAIDVLELVTSCSSNNVSVAHDASGNRIYYLT	2810
Qy	2807	PTTFLARAAWETARHTPIINSLNGNIIMYAPTLLWARMLTMHTFSIILIAQLEKALDCQI	2866
Db	2811	POVPLAKAAWETAKHSPVNSWLGNIIMYAPTLLWARIVLTMHTFSVLQSQSQLEKTLAFEM	2870
Qy	2867	YGACYSIEPLDLQPIIERLHGLSAPTLHYSYSGEINRVASCLRKGLVPPIURWHRARSV	2926
Db	2871	YGSVYSVTPLDLPAIIQRLHGLSAFSLHYSYSGEINRVASCLRKGLGVPPIURWHRARAV	2930
Qy	2927	RAXLLSGGGAATCGYLFNWARVTKLTPIPAASQLDLSGWFVAGYSGGDIYHSLISRA	2986
Db	2931	RAXLIIAQGGHAAICGLIYFNWA VKTKRLTPLADARLDLSSWFTVGAGGDIYHSMRSA	2990
Qy	2987	RPRWFFCLLLLSVGVGYILLPNR 3010	
Db	2991	RPRNLLCLLLLSVGVGYIFLLPAR 3014	

RESULT 10

QJ1303

genome polyprotein - hepatitis C virus (isolate HC-J6)

N:Contains: capsid protein C; envelope protein M; hepatitis C virus (isolate HC-J6) nonstructural protein NS5; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004

C:Accession: JQ1303

R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Kura, K.; Iizuka, H.; Machida, A.; Miyajima, T.; et al. 1991

A:Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a patient with chronic active hepatitis

A:Reference number: JQ1303; PMID:9204440; PMID:1658196

A:Accession: JQ1303

A:Molecule type: genomic RNA

A:Residues: 1-3033 <OK>	
A:Cross-references: UNIPROT:P26660; UNIPARC:UPI0000131E25; GB:D00944; NID:g221650; PIDN:	
A:Experimental source: isolate HC-J6 from a Japanese individual	
C:Superfamily: hepatitis C virus genome polyprotein	
C:Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transmembrane	
F:2-115/Product: capsid protein C #status predicted <CPC>	
F:116-191/Product: envelope protein M #status predicted <BPM>	
F:192-389/Product: major envelope protein E #status predicted <MEE>	
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>	
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>	
F:1011-1619/Product: hepatitis C virus nonstructural protein NS3 #status predicted <NS3>	
F:1316-1321/Region: nucleotide-binding motif B	
F:1320-1323/Region: DEXH motif	
F:1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>	
F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>	
F:2018-3033/Product: nonstructural protein NS5 #status predicted <N05>	
F:196,209,1234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,28	
Query Match 75.2%; Score 12036.5; DB 1; Length 3033;	
Best Local Similarity 71.9%; Pred. No. 0;	
Matches 2189; Conservative 338; Mismatches 472; Indels 45; Gaps 7;	
QY	1 MSTNPKPQRTKNTNRNRRPQDVFPGGQIVGGVYLLPRGRPLGVRAATKASERSQPRG 60
DB	1 MSTNPKPQRTKNTNRNRRPQDVFPGGQIVGGVYLLPRGRPLGVRAATKASERSQPRG 60
QY	61 RRQIPKARPEGRMAQPGPWPPLYNGEGLHAGWLLSPRSPSPSGPTDPRRSNGL 120
DB	61 RRQIPKARPEGRMAQPGPWPPLYNGEGLHAGWLLSPRSPSPSGPTDPRRSNGL 120
QY	121 KVIDTLTCGADLGMXYPLVAGPLGGARALAHGVRLVDGYNVATGNLPCSPSIFLLA 180
DB	121 KVIDTLTCGADLGMXYPLVAGPLGGARALAHGVRLVDGYNVATGNLPCSPSIFLLA 180
QY	181 LLSCLTIPASAYEVNRVSGIYHVNTDCSNSIVYEAADVIMHTPCGVCVQEGNSSRCW 240
DB	181 LLSCLTIPASAYEVNRVSGIYHVNTDCSNSIVYEAADVIMHTPCGVCVQEGNSSRCW 240
QY	241 ALTPTLAARNASVPTTIRRHVDLVLGTAAFCASAMYVDLGCISFLVSQLPFTSPRRHET 300
DB	241 PVSNVAVQPGALTOGLRTHIDMVMSATLCSALYVGLCGGVMLAAQMFIVSPQHWF 300
QY	301 VQDCNCSIIYCHVSGHRMAMDMANWSPPTALVYSQLLRTPQAVVDMVAGAHGVLAGLA 360
DB	301 VQDCNCSIIYCHVSGHRMAMDMANWSPPTALVYSQLLRTPQAVVDMVAGAHGVLAGLA 360
QY	361 YYSVMGVNWKVLIIVALLFAGVDGTHTTGRVAGHTSGFTSLFSSGASOKIQLVNTGWS 420
DB	361 YFSMQGAWAKVVVILLAAAGVDAQTHTVGGTAHARTLTGTFSLGARQKQLINTGWS 420
QY	421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCPERMASCRPIDWFAQGSPITY--TKPN 478
DB	421 HINRTALNCNDSLHTGFLASIFYTHSNSSGCPERMASCRPIEAFRVGMGALQYEDNVN 480
QY	479 SSDQRPCHWYAPRPGCVVPASQVCGPVYCTPSPVVVGTDRSGVPYISWGENETDML 538
DB	481 PEDMRPYCHWYPRQCGVVSASSVCGPVYCTPSPVVVGTDRSGVPYISWGENETDML 540
QY	539 LNNTRPPGCVNFGCTWNNSTGFTKCGGPPCNIIGVGNRT--LICPTDCFRKHPEATYK 596
DB	541 LNNTRPPGCVNFGCTWNNSTGFTKCGGPPCNIIRADFNASMDLLCPTDCFRKHPTTYIK 600
QY	597 CGSGPWLTPRCILDVYPRLMWYPCNTLNFISFKVRVYGVGEHRLNAACNWTGRCNLED 656
DB	601 CGSGPWLTPRCILDVYPRLMWYPCNTNYTIKIRVYGVGEHRLTAACNFTGRCNLED 660
QY	657 RDRSLSPLLLSTTSEWQILPCAFITLTPALSTGLIHLHQNIVDVQYLYGVGSFVSAFKW 716
DB	661 RDRSLSPLLLSTTSEWQILPCAFITLTPALSTGLIHLHQNIVDVQYLYGVGSFVSAFKW 720
QY	717 EYIILLFLLADARVACACILWMLLTAQAEALENILVNLNAASVAGAHGILSLFVFFCAW 776
DB	721 EYVILLFLLADARVACACILWMLLTAQAEALENILVNLNAASVAGAHGILSLFVFFCAW 780

QY	777 YIKGRAPGAAYAFYGVWPELILLALLPPRAYALDREMAASCGAVLVGLVFLTLSPYK 836
DB	781 YIKGRVVPATYSLTGLWSFGLLLALPQOAYAYDASVHGQIGALLVLLTLFTLTGYK 840
QY	837 VFLTRLIWLOVPIITRAEAHMVWVPLNVGRGDAIILLTCAVPELIDFIDTKLLAIL 896
DB	841 TLLSRFLWMLCYLTLUAEMVQWAPMQRVGRDGIINAWAIFCGVDFDITKWLAVL 900
QY	897 GPLMWLQAGITRVFYFVRAQGLIRACMLVRKVAAGHYVQVMFKLGAALTGYVYVNHUTPL 956
DB	901 GPAYLLKGAITRVFYFVRAHALRMCTMVRHLAGGRYQVMVLLALGRWTGYIYDHLTPM 960
QY	957 RDWAHAGLRDLAVAPVPPVFSAMETKVITWGAATAACDIIILGLPVSARGKEIFLGPAD 1016
DB	961 SDWAANGLRDLAVAVEPIIFSPMEKKVIWMAETAACGDILHGLPVSAIRLGREVLLPAD 1020
QY	1017 SLEQGWRLIAPITAYSQOTRGVLGCIITSLTRDKNOVEGEVQVWSTATQSFATCING 1076
DB	1021 GYTSKWSLLAPITAYAAQOTRGLLTIIVSWTGRDKTEQAGEIQVLTSTVTSQFLGTTISG 1080
QY	1077 VCVTVHAGSKTLAGPKGITQMYTNVDLDLVGWAQAPPGARSMTPCSCSSDLYLVTRH 1136
DB	1081 VLVTVHAGNKILASRGVPTQMYSSAEGDLVWPSPPGPKSLPCTCGAVDLYLVTRN 1140
QY	1137 ADVIPVRRRGRDSSRLSPVSYLKGSSGGPLLCPSGHVGVVFPRAAVCTRGVAKAVDFI 1196
DB	1141 ADVIPARRRDKRGALLSPRLSTLKGSSGGPVLCPRHAGVGVFPRAAVCSRGVAKSIDFI 1200
QY	1197 PVSMETMTMSPVFTDNSTPPAVPQTFQVAHLHAPTCGSKSTKVPAAAYAGYKVLVLP 1256
DB	1201 PVETLDIVTSPFSDNSTPPAVPQTVQVGLHAPTCGSKSTKVPAAAYAGYKVLVLP 1260
QY	1257 SVAATLFGAYMSKANGIDENIRTVTITGSGITYTYGKFLADGSCSGAYDIICD 1316
DB	1261 SVAATLFGAYLSKANGINIRTVTITGSGITYTYGKFLADGSCSGAYDIICD 1320
QY	1317 ECHSTDSTTILIGITVLDQAEATAGARLWLTATPPGSVTVPHNIEBIGLSNNGEIPFY 1376
DB	1321 ECHAVDSTTILIGITVLDQAEATAGARLWLTATPPGSVTVPHNIEBIGLSNNGEIPFY 1380
QY	1377 GKAIPIBAIKGGRHLIFCHSKKCKDELAALGTGLNNAVAYRGLDVSVPPIGDVVVA 1436
DB	1381 GRAIPLSYIKGGRHLIFCHSKKCKDELAALRGMLNAVAYRGLDVSVPITQGDVVVA 1440
QY	1437 TDALMTGFTGDFSDVDCNTCVTQTFPSLDPTTITTTTTPQDAVSRQRGRGGRS 1496
DB	1441 TDALMTGFTGDFSDVDCNVAVTQVDFSLDPTTITTTTTPQDAVSRQRGRGGRGL 1500
QY	1497 GIYRFVTPGPRPSGWFSSVLCVYDAGCAWVELTTPAETSRLRAYLNTPLGVPVQDHL 1556
DB	1501 GIYRVSTGERASGWFSSVLCVYDAGAAWVELTTPAETTVRLRAYFNTPLGVPVQDHL 1560
QY	1557 FWBSVFTGLTHIDAHFLSQTQAGDNFPYLVAYQATVCARQAQPPSPDWMMKCLIRLKP 1616
DB	1561 FWEAVFTGLTHIDAHFLSQTQSGENFAYLTAYQATVCARAKAPPPSPDWMMKCLIRLKP 1620
QY	1617 TLHGFTPLLYRLGAVONEVILTHPTIKYIMACNSADLEVVTSTWLVGVGLAALAYCLT 1676
DB	1621 TLVGFTPLLYRLGAVONEVILTHPTIKYIATCMQADLEVMTSTWLVAGVLAALAYCLT 1680
QY	1677 TGSVVVIGRIILSKPAPVDPREVLYQFDEMEBCASQLPIEQGMQLAEQFKQALGIL 1736
DB	1681 TGCVCIIIRLHVNRQAVVAPDKVEVLYEAFDEMEBCASRAALIEGRIETAEMLKSIQGLL 1740
QY	1737 QTATKQAEAAAPVSKWRALETFWAKHWNFIISGIQYLAGISTLPGNPAIASLMAFTAS 1796
DB	1741 QQASKAQODIQPAVOASWPKEQFWAKHWNFIISGIQYLAGISTLPGNPAVASMAFSA 1800
QY	1797 ITSPLTNTQNTLLENILGGMVAAQAPSAASAFVAGIAGAAVSGISGLGKVLVDILAGY 1856
DB	1801 LTSPLSTTTILLNITLGLWLASQIAPPAGATGVVSGLVGAAGVSGISGLGKVLVDILAGY 1860

QY	1857	AGVAGALVAFKMSGEVSTEDLVNLLPAILSPGALVGVGVCAAILRRHVHGEGEQAQVM	1916
DB	1861	AGISGALVAFKIMSGEKSPMEDVNVLLPGIILSPGALVGVCAAILRRHVHGEGEQAQVM	1920
QY	1917	NRLIAFASRGNHVSPTHVYPESDAAARVTOILSLTITQLKRLHWINEDCSTPCSGSW	1976
DB	1921	NRLIAFASRGNHVAPTHVYTESDAQRVTOQLLSLITISLRLRLHWNITDCPCPCSGSW	1980
QY	1977	LRDVMWDVICTVLTDFTKWLQSKLLPRLPGVPPFLSCQGYKGVWRGDGIMQITPCPGAQITA	2036
DB	1981	LRDVMWDVICTILTDFKNWLTSLKFLPMFGLPIFSCQKGYKGVWAGTGIMTTRCPGANIS	2040
QY	2037	GHVKNGSMRIIVGRTCSNTHGTFPINAYTTGCTPSPAPNYSRALRWRAAEYVEVTRV	2096
DB	2041	GNVRLSGSMRITGPKTCMNIQGTTFPINCYTEQCQVPKPAPNFKIAIWRVAASEYAEVTQH	2100
QY	2097	GDFHYVTGTTDNVKPCQVPAPEPTEVDGVRHLRYAPACKPLIREDTVFQVGLNQVIL	2156
DB	2101	GSYHYITGLTDNLKVPQCLPSPFPFSSVDGVQIIRFAPIPKPFPRDEVSFCVGLNSFVV	2160
QY	2157	GSQLPCEPEPDVTVLTSMLTDSHITAETAKRLARGSPPLASSASOLSAPSLKATCT	2216
DB	2161	GSQLPCEPEPDVTVLTSMLTDSHITAETAAARLARGSPPSASSASOLSAPSLRAICT	2220
QY	2217	THHDSPADLIEANLLWRQMGCGNITRVESENKVVILDSFELHAEGDEREISVAAEILR	2276
DB	2221	THGKAYDVMVDANLF---MGSDVTRIESESKVVVLDLSDPMEERSDLEPSIPSEYML	2276
QY	2277	KSRKPSALPIWARDYNNPPLLESWKDDPYVPVHVGCPDLPPTKAPPTPPPRKRTVULT	2336
DB	2277	PKKREPPALPAWARDYNNPPLVESWKRPDYQATVAGCALPPPKKPTPTPPRRRTVGLS	2336
QY	2337	ESNVSSALAEATKTFG-----SSGSADVSGTATLPDLASDDGDKGSVDESYS	2386
DB	2337	ESSIADALQQLAKFGQPPSGDSGLSTGADAAGSGRTPPDEIAL-----STGSLIS	2390
QY	2387	SNPPLEGEFGDPDL-----SDGSMSTVSEASESDVWCSSMSYWT	2426
DB	2391	SMPLEGEFGDPLRPEQVELQPQGGVVTPGSGSGSWSTCSEB-DDDSVWCSSMSYWT	2449
QY	2427	GALITPCAABESKLPINPLNSLLRHNNVYATTSRASLRQKVTFRQLQVLDLDDHYRDV	2486
DB	2450	GALITPCSPPEEKLPINPLNSLLRYHNKYCTTSSASLRKAKVTFRDMQALDAHYDSV	2509
QY	2487	LKEMAKASTVAKALLSTEEACKLPHPHSAKFGYGAKDVRNLSSRAVNHRSVYEDLL	2546
DB	2510	LKDILKASKVTARLLTLEAQLTPPHSARSKYGFGAKEVRSLSGRAVNHKSVMKDLL	2569
QY	2547	EDTETPIDTTIMAKSEVFCVQPEKGRKPARLIVFPDLGVRVCEKMALYDVVSTLPOAVM	2606
DB	2570	EDTQTPITPTIMAKNEVFCVDPFKGKKAARLIVPDLGVRVCEKMALYDIITQKLPOAVM	2629
QY	2607	GSYGFQYSPKQRFVLVNTWKSCKPMGFSYDTRCFDSTVTTESDIRVEESYQCCDLAP	2666
DB	2630	GASYGFQYSPAQRVEFLKAWAEKKDPMGFSYDTRCFDSTVTTERDIRTEESYRACSLPE	2689
QY	2667	BARQAIRSLTERLYTGGPLTNSKGCNGYRRCCRASGVLTTSGNTLTCTVLKATAACRAK	2726
DB	2690	EAHTAIHSUTERLYYGVGPMFNKSGTCGYRRCCRASGVLTTSMGNTITCTVKKALACKAAG	2749
QY	2727	LODCTMLVNGDDLWIVICESAGTQDEAAALRAFTEAMTRYSAAPGDPPOPEYDLELITSCS	2786
DB	2750	IATPMLVCGDDLWIVICESQGTEDERNLRAFTEAMTRYSAAPGDPPEYDLELITSCS	2809
QY	2787	SNVSVANDASGRVYVLTDPPTPLARAAWETARHTPPINSWLGINIIMYAPTIAWARMILMT	2846
DB	2810	SNVSVALGPQGRVYVLTDPPTPLARAAWETVRHSPVNSWLGNITIQYAPTIAWARMVLT	2869
QY	2847	HFESILLAQOEKALDCQIYGACYSIEPLDLPQIIRLHGLSAFTLHSHYSPEGEINRVAS	2906
DB	2870	HFESILMAQDTLDQNLNFEMYGAVSVSPDLUPAIIERLHGLDAFSLHTYTTTTPHELTRVAS	2929
QY	2907	CLURKLGVPPLTRWRHARSVRAKLIASQCGRAATCGRYLFWAVRTKLTPTIPAAASQLDL	2966

DB	2930	ALRKUGAPPLRAWKSRAVRAVSLISRGGRAAVCGRYLFWNAVKTKLKTPLPEARLLDL	2989
QY	2967	SGWFVAGYSGGDIYHSLSRARPRWFPPLCLLLLSVGVGYVLLPNR	3010
DB	2990	SSWFTVAGGGDIYHSVSARPRLLLLGLLLLFVGVGLFLLPAR	3033
RESULT 11			
GNWVJ8			
genome polypotein - hepatitis C virus (strain HC-J8)			
N:Contains: capsid protein C; envelope protein M; hepatitisvirin (EC 3.4.21.98) (nonstructuc			
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5			
C:Species: hepatitis C virus			
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004			
A:Accession: A40250; PQ0397; PQ0559			
R:Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Linuka, H.; Tanaka, T.; Fukuda, S.;			
Virology 188, 331-341, 1992			
A:Title: Full-length sequence of a hepatitis C virus genome having poor homology to repo			
A:Reference number: A40250; MUID:92230232; PMID:1314459			
A:Molecule type: genomic RNA			
A:Residues: 1-3033 <OKA>			
A:Cross-references: UNIPROT:P26661; UNIPARC:UPI0000131E27; GB:D10988; GB:D01221; NID:g22J			
J.Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.L.			
J. Gen. Virol. 73, 1131-1141, 1992			
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e			
A:Reference number: PQ0393; MUID:92268871; PMID:1316939			
A:Accession: PQ0397			
A:Molecule type: genomic RNA			
A:Residues: 2678-2754 <CHA>			
A:Cross-references: UNIPARC:UPI0000174A02; DDBJ:D10134			
R:Kato, N.; Ootsuvara, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno,			
Biochem. Biophys. Res. Commun. 181, 279-285, 1991			
A:Title: Distribution of plural HCV types in Japan.			
A:Reference number: PQ0554; MUID:92068204; PMID:1720309			
A:Accession: PQ0559			
A:Molecule type: mRNA			
A:Residues: 2678-2729 <KAT>			
A:Cross-references: UNIPARC:UPI00000F5263; GB:D10562; GB:D90518; NID:g221523; PIDN:BAA01			
C:Superfamily: hepatitis C virus genome polypotein			
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural			
F:1-115/Product: capsid protein C #status predicted <CPC>			
F:116-191/Product: envelope protein M #status predicted <EPM>			
F:192-389/Product: major envelope protein E #status predicted <MEE>			
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>			
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>			
F:1011-1619/Product: hepatitisvirin #status predicted <NS3>			
F:1234-1241/Region: nucleotide-binding motif A (P-loop)			
F:1316-1321/Region: nucleotide-binding motif B			
F:1320-1323/Region: DEXH motif			
F:1620-1866/Product: nonstructural protein NS4a #status predicted <N4a>			
F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4b>			
F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>			
F:196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,23			
Query Match 74.8%; Score 11972.5; DB 1; Length 3033;			
Best Local Similarity 70.9%; Pred. No. 0;			
Matches 2158; Conservative 369; Mismatches 473; Indels 43; Gaps 9;			
QY	1	MSTNPKPQKTKRNTNRPRQDVKFFGGGQIVGGVLLPRRGLGVRAATRKASERSQPRG	60
DB	1	MSTNPKPQKTKRNTNRPRQDVKFFGGGQIVGGVLLPRRGLGVRAATRKASERSQPRG	60
QY	61	RQPIPKARRPEGRAWAQPGYWPVLYGNEGLGWAGLLSPRGSRPSWGPTDPRRRRNIG	120
DB	61	RQPIPKARRSTGKSWGKPGYWPVLYGNEGCGWAGLLSPRGSRPTGWTDPDRHRSNIG	120
QY	121	KVIDTLTCGFADLMGVIPLVGAPLGGARALAHGVRLVDGNYATCNLPGCSFSIFLLA	180
DB	121	RVIDTITCGFADLMGYIPVVGAPVGVARALAHGVRLVDGNYATCNLPGCSFSIFLLA	180

181 LLSCLTIPASAYEVRNVSGIYHVNTDCSNSSIVYEAAADVIMHTPGVPCVQBOGNSRCWV 240
181 LLSCVTPVSAVERNISSYYATNDCSNSSITWLTDAVLHLPVCGPCENDNGTILHWI 240
241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCASAMVUGDLCSIFLVLSOLFTEPSRHHET 300
241 QVTPTNAVVRHGRGALTRSLRTHVMDI VMAATAFCSALYVGDVCGAVMILISOAFVSPQRHNF 300
301 VQDCNCSIYPCGHVSRHMAWMDMMNSPTTALVVSOLLRIPOAVVDMVAGAHGVLAGLA 360
301 TQECNCSIYQCHITGRHMAWMDMLNSPTLTMILAYAAVPELVLEIIFGHHGVVFGLA 360
361 YYSMVGNWAKVLIVALLFAGVGETHTTGRVAGHTTSGFTSLFSSGASQIKLIVNTNGSW 420
361 YFSMOGAWAKVIAILLVAGVDATYSSGOEAGRTVAGFAGLTGTAKQNLXYLINTNGSW 420
421 HINRTALNCNDSLOTGFFAALFVAHKFNSSGCPERMASCRPIDFAQGWGPIY - TKPN 478
421 HINRTALNCNDSLOTGLASLFYTHKFNSGCPERLSSCRGLDDFRIGWGTLEYETNVTN 480
479 SSDORPVCMHVAPRBCGVVPASQVCGPVCTPSPVVGTTDRSGVPTYSGWNETDVML 538
481 DGMRPVCMHYPRPCGIVPARTVCGPVICTPSPVVVGTTDKQGVPTYWGENETDVL 540
539 LNNTRPQGNWFGCTWMNSTGFTKTCGPPCNTIGGVGNRT--LICPTDCFRKHPEATYTK 596
541 LNSTRPRGANFGCTWNGTFTKTCGAPCRIRKDYNSIIDLLCPTDCFRKHPEATYTK 600
597 CGSGPWLTPRCLVDYVYRLMHPCTLNFSIPKVMYVGGVEHRLNAAACNTRGRRCNLED 656
601 CGAGPWLTPRCLVDYVYRLMHPCTVNTFTFKARMYVGGVEHRSAAACNTRGRDCRLED 660
657 RDRSELSPILLSSTEWOLPCAFITLPAISLGLIHLHQNIVDVQYLVGVGSFAKFW 716
661 RDRGQQSPILLSSTEWAVLPCSPSDDLPAISLGLIHLHQNIVDVQYLVGLSPALTRYIVKW 720
717 EYILLFLLADARVCAACLMWMLIAOAEAALENLVNNAASVAGAGHILSFLVFFCAAW 776
721 EWVILLFLLADARICACLMWMLIILGQAEAALEKLIILHSAASANGPLWFFIFFTAAW 780
777 YIKGRAPGAAFYGVMPILLLLALPPRAYALDREMAASCGGAVILVGLVFLTSPYYK 836
781 YLKGVRVPVATYSVLGLWSFLLVLAUPOQAYALDAEAGBELGAILVIIISFTLPAYK 840
837 VFLTRLIWLQYFTRAHAHQVVPVPLNVGRGDAIILLTCAVHPELIPDITKLLAIL 896
841 ILLSRVWMLSYMLVLAEAQIQWVPVLEVRGGRDGIIVWAVILHPLRVFVTKWLLAIL 900
897 GPLMVLOAGITRVPYFVRAOGLIRACMLVRKVAGHYVQVFMKLGALTGTYYVYNHLTPL 956
901 GPAYLLKASLRIPYFVRAHALLVCTLVKHLAGARYIOMLLITIGRWGTGYIYDHLSP 960
957 RDWAHAGLRDLAVAPVVFSAEMETKVTWCADTAACGDIILGLPVSARRKEIFLGPAD 1016
961 STWAAQGLRLDAIAVEPVVFPMKEKVIWGAETVACDIIHLGLPVSARLREVLLGPAD 1020
1017 SLEGOWRLLAPITAYSQOTRGVLGCIITSTIGRDKNOVEGEVQVSVSTATQSFLATCING 1076
1021 GYTSKGWKLAPITAYTQOTRGLLGAIVVSLTGRDKNEQAGQVQLSSVTTQTLGTSISG 1080
1077 VCWTVYHAGSKTLAGPKGPIQTMYNVNVDLIVGHOAPPGARSMTPCSGSSDLVLVTRH 1136
1081 VLWTVYHAGNKTLAGPKGPVTQMYTSAEGLVGMPSPPGTSKSLDPCTCGAVDVLVTRN 1140
1137 ADVIPVRRGRSGSLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAAYCTRGVAKADVFI 1196
1141 ADVIPVRRKDRRGALLSPRLSTLKGSSGGVPLCSRGHAGVGLFRAAAYCAVAKSIDFI 1200
1197 PVSEMETTMRSPVFTNDSTPPAVPOTFOVHLHAPTSGKSTKVPAAVAAQYKVLVNLN 1256
1201 PVESLDVATRTPSFSDNSTPPAVPOSQVGYLHAPTSGKSTKVPAAVAAQYKVLVNLN 1260
1257 SVAATLFGAYMSKAHGDPIRNTGVRTITTTGSGITYSTYKFLADGGCGGAYDIICD 1316

1261 SVAATLFGAYMSKAHGINPNIIRNTGVRTITTTGSDIYIYKFIADGGCAAGAYDIICD 1320
1317 ECHSTDSTTILGIGTVDQAETAGARLVVLATATPPGSVTVPHPNIEEIGLSNNGEIPFY 1376
1321 ECHSVDAITLIGTVDQAETAGARLVVLATATPGTTPPHSNIEEVALGHEGEIPFY 1380
1377 GKAIPEALKGGHLLIFCHSKKCKDELAALTGCLGNAYAYRGLDVSVPPTPGDVVVVA 1436
1381 GKAIPLAFIKGGHLLIFCHSKKCKDELAALRGMGNAYAYRGLDVSVPPTQGDVVVVA 1440
1437 TDAIMTGTCDPDSVDIDCNTCTQTVDPSLDPTFTTETTTVPQDASRSRRGRTGRGRS 1496
1441 TDAIMTGYTGDFOVDIDCNVAVSQIVDFSLDPTFTTITQIVPQDASRSRRGRTGRGL 1500
1497 GIYRFTVTPGSPGMPDSSVLCYDAGCAWYELTTPAETSVRLRAYLNTPLGVPDCHLE 1556
1501 GVTRYVSSGERPSPGMPDSSVLCYDAGCAWYELTTPAETTVRLRAYLNTPLGVPDCHLE 1560
1557 FWESVFTGLTHIDAHFLSOTKQAGDNFPYLVAVQATVCARAQAPPPSPSDOMKWLIRLKP 1616
1561 FWEAVFTGLTHIDAHFLSOTKQAGDNFPYLVAVQATVCARAQAPPPSPSDVMWKLIRLKP 1620
1617 TLHGPTPLLYRLGAVQNEVILTHPIKYIMACHSADLEVTTSTWLVGGVLAALAAAYCLT 1676
1621 TLGTPTPLLYRLGAVTNEVILTHPIKYIATCMQADLEIMTSSWVLGAGVLAALAAAYCLA 1680
1677 TGSVIVIGRIILSKPAAVDPREVLYOEFDEMECASQLPVIEQGMOLAEQFKQKALGLL 1736
1681 TGCISIIGRUHLNDRVVVVADPKELLYEAFDEMECASQALIEBEGQMAEMLKSKIQGLL 1740
1737 QTATKQAEAAAAPVSVESKWRALFTWAKHMNFISGIIYLAGLSTLPGNPAIASLMAFTAS 1796
1741 QOATRQADITQALQSSWPKLEQFWAKHMNFISGIIYLAGLSTLPGNPAVASMAFSA 1800
1797 ITSPLTQNTLLFNILGGWAAQLAPPSAASAFVAGIAGNAGSIGLCKVLVDILAGY 1856
1801 LTSPLTSTILLINIMGVLASQIAPPAGATGFVVSGLVGAAGSIGLGLKTLVDVLAGY 1860
1857 AGVAGALVAFKWSGEVPSSTEDLVNLLPALISPCALVGVVCAAILRRHVGGGAGVOM 1916
1861 AGISGALVAFKWSGEVPSSTEDLVNLLPALISPCALVGVVCAAILRRHVGGGAGVOM 1920
1917 NRLIAPASRGNHVSPTHVYVESDAAARVQIILSLTITQLIKRLHQMINEDCSTPCSGW 1976
1921 NRLIAPASRGNHVAPTHVYVESDASQRTQVLSLTITSLRLRHAMITECPCVPCSGW 1980
1977 LRVDWDWICTVLTDFTKTLQSLKLLPRLPGVPFLSCQYKGVWGRGDGIMQTTCPGQAIA 2036
1981 LODIWDWVCSLITDFTKWLSSKLLPKMGPFFISQYKGVWAGTGVMTTRCPCGANIS 2040
2037 GHVKNXGSMRIVGPRTCSNTHGTPINAYTTGCTPSPAPNYSRALMRVAABEYEVTRV 2096
2041 GHVRMGTKITGPKTCLNMQGTFPINCYTEGCPVKPPPNYKTAIRMRVAABEYEVTRV 2100
2097 GDFHYVTGTTDVKCPQAPPEFFTEVDGVRHRYAPACKPDLREDDVTQVGLNQYLV 2156
2101 GSFSYVTLTSDNLKVPQVPAPEFFSWDVGQIHRFAPVPFPFRDEVTFTVGLNSFV 2160
2157 GSQPCPEPDPVTVLTSMITDPSHTAETAKRRLARGSPPSLASSSSASQLSAPLSKATCT 2216
2161 GSQPCPEPDPTEVLASMLTDPSHITAEAAARLARGSPPSQASSSSASQLSAPLSKATCT 2220
2217 THHSDPADLITAEALLMRQEMGNITRVESKNVILDSFPLHABGEREISVAEILR 2276
2221 THKTAYDCMDVANDL.F---MGGDVTRIESDSKVIILSDLSMTSVEDDRSPSVSEVLI 2276
2277 KSRKPPSALPTIWARPDYNNPPLLESKWDDYVPPVVGCPFLPPTKAPPTPPRRKRTVILT 2336
2277 KRRKPPPALPPWARPDPYNPVLIETWKRPFGYBPPTVLGALPPTPTQTPVPPRRRRKAVLT 2336
2337 ESNVSSALAEALATKTFG-----SSGSSAVDSGTATALPOLASDDGDK--GSDVESYSS 2387

Db	2337	QDNVEGVLR	EMADKVLSP	LDNNDSGH	STGADTG-----	GDIVQQPS	DDTAASEAG	SS	2391		
Qy	2388	MPLEGE	GGDPDL-----	SDG-----	SWSTVSE	EEASED	VVCCSM	SYTW	2427		
Db	2392	MPLEGE	GGDPDL	EPFVGS	APPSEGE	CEVIDS	SKSWTS	DSQ-EDSV	ICCSMSY	2450	
Qy	2428	ALITPCA	EEESKLP	INPLNS	LLRHNM	VYATTS	RSASLR	QKVKV	TFDR	LQVLDH	2487
Db	2451	ALITPC	GEPEEL	PIPLNS	LMRFHN	KVYTS	TSRSAS	LRACKV	TFDR	VQVLDH	2510
Qy	2488	KEMKAS	TVKAKL	ISIEAC	KLTPH	SAKSKF	YGAKD	VRLNS	SRAN	HIRS	2547
Db	2511	QDKRA	ASAR	LVVEAC	ALTTPH	SAKSYG	FGAKE	VRSL	RRAN	HIRS	2570
Qy	2548	DTETP	IDTTIM	AKSEV	FCVQ	PEKGR	KPARL	IVFP	DGLV	RVC	2607
Db	2571	DQHTP	IDTTIM	AKNEV	FCIDP	TGKGK	PARL	IVPD	LGRV	RVC	2630
Qy	2608	SSYGF	OYSPK	QVRE	FLVNT	WKS	KCPMG	SFYD	TRCF	DSVT	2667
Db	2631	PSYGF	OYSPA	ERVDF	LLKAWG	SKDPM	GFSDY	TRCF	DSVT	TERD	2690
Qy	2668	ARQAR	ISRL	TERLY	IG	PLTNS	GQCGV	RRCAS	GVLT	TS	2727
Db	2691	ARTV	HSRL	TERLY	VG	PMTNS	GQCGY	RRCAS	GVLT	TS	2750
Qy	2728	QDCTM	LVNGD	DLV	ICES	AGTQ	EDAA	ALRAF	TEAM	TRY	2787
Db	2751	VDPM	LVCGD	DLV	ISES	QNE	DEERN	LRAF	TEAM	TRY	2810
Qy	2788	NVSVA	HDAS	GKRV	YVLT	TRDPT	TLARA	AWET	ARHT	PINS	2847
Db	2811	NVSVA	LD	SGRR	RYFL	TRDPT	PI	TRA	AWET	VRHS	2870
Qy	2848	FFSILL	AQBLE	KALD	COIY	GACYS	IEPL	DLQ	LIER	LHGL	2907
Db	2871	FFSILL	AQD	TLN	QNF	YGVAV	SYN	PLD	PAI	IERL	2930
Qy	2908	LRLG	VPP	LRT	WRH	ARS	VRAK	L	SGG	RAAT	2967
Db	2931	LRLG	GA	PP	L	AWK	SRAR	VRAS	L	IAQ	2990
Qy	2968	GMFV	AGY	GGD	IYH	LSR	APR	WFPL	CLLL	SVG	3010
Db	2991	GMFT	VAG	GGD	IYH	SVH	SHAR	PRLL	LL	CLLL	3033

RESULT 12

S18031

genome polypeptide - hepatitis C virus (isolate JK2) (fragment)

N;Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein

C;Species: hepatitis C virus

A;Variety: isolate JK2

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.

submitted to the EMBL Data Library, September 1991

A;Description: Sequence analysis of putative structural regions of Hepatitis C Virus isolate JK2

A;Reference number: S18029

A;Accession: S18031

A;Molecule type: genomic RNA

A;Residues: 1-782 <HON>

A;Cross-references: UNIPROT:Q69950; UNIPARC:UPI0000178520; EMBL:X61593

A;Experimental source: isolate JK2

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural F1;191/Product: core protein #status predicted <MAT1>

F;192-383/Product: envelope protein 1 #status predicted <MAT2>

F;384-733/Product: NS1/E2 protein #status predicted <MAT3>

F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 24.8%; Score 3977; DB 2; Length 782;

Best Local Similarity 91.9%; Pred. No. 4.6e-233;

RESULT 12

S18031
genome polyprotein - hepatitis C virus (isolate JK2) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK2
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus isolate
A:Reference number: S18029
A:Accession: S18031
A:Molecule type: genomic RNA
A:Residues: 1-782 <HON>
A:Cross-references: UNIPROT:O68950; UNIPARC:UPI0000178520; EMBL:X61593
A:Experimental source: isolate JK2
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F:1-191/Product: core protein #status predicted <MAT1>
F:192-383/Product: envelope protein 1 #status predicted <MAT2>
F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 24.8%; Score 3977; DB 2; Length 782;

Best Local Similarity 91.9%; Pred. No. 4.6e-233;

Matches	719;	Conservative	27;	Mismatches	36;	Indels	0;	Gaps	0;							
Qy	1	MSTNPKPQRTKXNTNRR	PQDKV	FGGGQIVGGVYLL	PRRGPR	PLGVRA	TRKASERSQ	PRG	60							
Db	1	MSTNPKPQRTKXNTNRR	PQDKV	FGGGQIVGGVYLL	PRRGPR	PLGVRA	TRKTSERSQ	PRG	60							
Qy	61	RRQPIKARRPEGRAWAQ	PGY	PWPLYGNEGLWAGWLL	SPGRSP	SGWPTD	PRRRSRNLG	120								
Db	61	RRQPIKARQPEGRAWAQ	PGY	PWPLYGNEGLWAGWLL	SPGRSP	SGWPTD	PRRRSRNLG	120								
Qy	121	KVIDTLTCGFADLMGYI	PLVGA	PLGAARALAHGVRL	VEDG	VNYATCN	PGCSFS	FLLA	180							
Db	121	KVIDTLTCGFADLMGYI	PLVGA	PLGAARALAHGVRL	VEDG	VNYATCN	PGCSFS	FLLA	180							
Qy	181	LLSCLTI PASAYEV	RNVSGIYHV	TNDCSN	SIYVEAAD	IMHTPG	CVPCVOEG	SSRCWV	240							
Db	181	LLSCLTI PASAYEV	RNVSGIYHV	TNDCSN	SIYVEAAD	IMHTPG	CVPCVLEGN	SSRCWV	240							
Qy	241	ALTPTTLAARNASV	PTTTIRRH	VDLLVGTAA	FCSAMV	VGDL	CGSIFLVS	QLFTFSP	RRHET	300						
Db	241	ALTPTTLAARNSS	PTTTIRRH	VDLLVGA	AFCSAMV	VGDL	CGSVFLVS	QLFTFSP	RRYET	300						
Qy	301	VQDCNCSLYPGHVS	GRMAWDM	MMNWS	PTTALV	VSQ	LLRIPOAV	VDVMA	GAHGV	LAGLA	360					
Db	301	VQDCNCSLYSGHVS	GRMAWDM	MMNWS	PTTALV	VSQ	LLRIPOAV	VDVMA	GAHGV	LAGLA	360					
Qy	361	YYSWGNWAKVLI	VALLFAG	VDGETH	TTCRV	AGHTT	SGTSLF	SSCAS	QIKOL	LVNTG	SW	420				
Db	361	YYSWGNWAKVLI	VMLLFAG	VDGRTQ	VTGAQ	AGHTT	SGLAS	LFTPG	SQIKQL	VNSNG	SR	420				
Qy	421	HINRTALNCNDS	LIOTGFFA	ALFYAH	KFNSS	SGC	PERMAS	CRPID	WFAQ	GWGP	ITYTK	PNSS	480			
Db	421	HINRTALNCNDS	LIKTGFLA	LFYTH	KFNA	AGC	PERMAS	CCSID	TDPQ	GWGP	ITHVV	VPNST	480			
Qy	481	DQRPYCHWYAP	RCGPVVP	ASQ	CGVPV	CTP	SPV	VVGT	TRD	SGVP	PTYS	GENE	TDV	MLN	540	
Db	481	DQRPYCHWYAP	RCGPVVP	ASQ	CGVPV	CTP	SPV	VVGT	TRD	FGAP	PTYN	GANE	TDV	VLN	540	
Qy	541	NTRPPQGNWFG	CTW	MNSTG	FTKTC	CGG	PCNIGG	VNRT	LICPT	DC	FRK	HEAT	YTK	CGS	600	
Db	541	NTRPPQGNWFG	CTW	MNSTG	FTKTC	CGP	PCNIGG	VNNT	LITC	PTDC	FRK	HEAT	YTK	CGS	600	
Qy	601	PMLTPRCLVD	YPYRLWH	YPTCLN	FSIF	KVMY	VGVG	VEHRL	NAAC	NW	TRG	ERC	NLE	DRS	660	
Db	601	PMLTPRCMV	HYPYRLWH	YPTCVN	FTVK	VMY	VGVG	VEHRL	TAA	CN	W	TRG	ERC	NLE	DRS	660
Qy	661	ELSPLLSTTE	WQILPCA	FTTLP	ALST	GLIH	QNI	VDVQ	YLV	GVGS	AFV	SFA	KW	EYIL	720	
Db	661	ELSPLLSTTE	WQILPCS	FTTLP	ALST	GLIH	QNI	VDVQ	YLV	GVGS	AFV	SFA	KW	EYIL	720	
Qy	721	LFLPLLADAR	VCACILMM	LLIAQ	AEAL	ENLV	LNA	ASVAG	AHG	ILSL	FLV	FFCA	AWY	IKG	780	
Db	721	LFLPLLADAR	VCACILMT	LLIAQ	AEAL	ENLV	LNA	ASVAG	AHG	ILPS	FLV	FFCA	AWY	IKG	780	
Qy	781	RL	782													
Db	781	RL	782													

RESULT 13
S19876
genome polyprotein - hepatitis C virus (isolate JK5) (fragment)
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C;Species: hepatitis C virus
A;Variety: isolate JK5
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S19876
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A;Description: Sequence analysis of putative structural regions of Hepatitis C virus
A;Reference number: S18029
A;Accession: S19876
A;Molecule type: genomic RNA

RESULT 13

S19876
genome polyprotein - hepatitis C virus (isolate JK5) (fragment)
N:Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C:Species: hepatitis C virus
A:Variety: isolate JK5
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus isolate
A:Reference number: S18029
A:Accession: S19876
A:Molecule type: genomic RNA

A;Residues: 1-782 <HON>
A;Cross-references: UNIPROT:Q68953; UNIPARC:UPI00000EC530; EMBL:X61595; NID:g59486; PIDN
A;Experimental source: isolate JK5
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F;1-191/Product: core protein #status predicted <MAT1>
F;192-383/Product: envelope protein 1 #status predicted <MAT2>
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match		24.7%;	Score 3952;	DB 2;	Length 782;
Best Local Similarity		90.9%;	Pred. No. 1.5e-231;	Indels 0;	Gaps 0;
Matches 711;		Conservative 31;	Mismatches 40;		
Qy	1	MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRGRPLGVRAATKASERSQPRG	60		
Db	1	MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRGRPLGVRAATKASERSQPRG	60		
Qy	61	RRQPIPKARPEGRAMAQPGYPWPLYGNEGLWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120		
Db	61	RRQPIPKARPEGRAMAQPGYPWPLYGNEGLWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120		
Qy	121	KVTDLTTCGADLMGYIPLVGAPLGGAAARALAHGVRVLEGDVNVATGNLPGCSFSIFLLA	180		
Db	121	KVTDLTTCGADLMGYIPLVGAPLGGAAARALAHGVRVLEGDVNVATGNLPGCSFSIFLLA	180		
Qy	181	LLSLCTIPASAYEVRNVSGIYHVNTDCSNSISIVYEAADVMHTPGCVPCVOEGNSSRCWV	240		
Db	181	LLSLCTIPASAYEVRNVSGIYHVNTDCSNSISIVYEAADVMHTPGCVPCVOEGNSSRCWV	240		
Qy	241	ALPTTLAARNASVPTTIRRHVDLLVGTAAAFCSAMYVDGLCGSIFLVSQLFTSPRRHET	300		
Db	241	ALPTTLAARNASVPTTIRRHVDLLVGTAAAFCSAMYVDGLCGSIFLVSQLFTSPRRHET	300		
Qy	301	VQDNCISYIPGHVSGHRMAMDMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLGAGLA	360		
Db	301	VQDNCISYIPGHVSGHRMAMDMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLGAGLA	360		
Qy	361	YYSMVGNWAKVLIIVALLFAGVDGTHTRVAGHTTSGTSLFSSGASQKIQLVNTNGSW	420		
Db	361	YYSMVGNWAKVLIIVALLFAGVDGTHTRVAGHTTSGTSLFSSGASQKIQLVNTNGSW	420		
Qy	421	HINRTALNCNDSLOTGFFAALFYAHKFNSSCCPERMASCRIPIIDWFAQGWGPIYTKPNSS	480		
Db	421	HINRTALNCNDSLOTGFFAALFYAHKFNSSCCPERMASCRIPIIDWFAQGWGPIYTKPNSS	480		
Qy	481	DQPYCHYAPRPGVVPASQVCGPVYCTFSPVVGTTDRSGVPTYSWGENETDVLIN	540		
Db	481	DQPYCHYAPRPGVVPASQVCGPVYCTFSPVVGTTDRSGVPTYSWGENETDVLIN	540		
Qy	541	NTRPQGNWFCGTWNSTGFTKTCGGPPCNIGGVNRTLICTPTDCFRKHPEATYTKCGSG	600		
Db	541	NTRPQGNWFCGTWNSTGFTKTCGGPPCNIGGVNRTLICTPTDCFRKHPEATYTKCGSG	600		
Qy	601	PWLTPRCIVDYPYRLWHYPCTINFSIFKVRMYVGVGVEHRLNAACNWTGERCNLEDRDS	660		
Db	601	PWLTPRCIVDYPYRLWHYPCTINFSIFKVRMYVGVGVEHRLNAACNWTGERCNLEDRDS	660		
Qy	661	ELSPILLSTTEWQILPCAFPTTLPALSTGLIHLHQNIVDVQVLYGVGSFAIKWEYIL	720		
Db	661	ELSPILLSTTEWQILPCAFPTTLPALSTGLIHLHQNIVDVQVLYGVGSFAIKWEYIL	720		
Qy	721	LLFLLADARVCACILMMWLLTAQEAALENVLNAASVAGAHGILSLFVFPFAAWYIKG	780		
Db	721	LLFLLADARVCACILMMWLLTAQEAALENVLNAASVAGAHGILSLFVFPFAAWYIKG	780		
Qy	781	RL 782			
Db	781	RL 782			

genome polyprotein - hepatitis C virus (isolate JK4) (fragment)
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
C;Species: hepatitis C virus
A;Variety: isolate JK4
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
R;Accession: SI8032
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A;Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso
A;Reference number: SI8029
A;Accession: SI8032
A;Molecule type: Genomic RNA
A;Residues: 1-782 <HON>
A;Cross-references: UNIPROT:Q68952; UNIPARC:UPI0000178521; EMBL:X61594
A;Experimental source: isolate JK4
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F;1-191/Product: core protein #status predicted <MAT1>
F;192-383/Product: envelope protein 1 #status predicted <MAT2>
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match		24.6%;	Score 3942;	DB 2;	Length 782;
Best Local Similarity		90.8%;	Pred. No. 6.1e-231;	Indels 0;	Gaps 0;
Matches 710;		Conservative 31;	Mismatches 41;		
Qy	1	MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRGRPLGVRAATKASERSQPRG	60		
Db	1	MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRGRPLGVRAATKASERSQPRG	60		
Qy	61	RRQPIPKARPEGRAMAQPGYPWPLYGNEGLWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120		
Db	61	RRQPIPKARPEGRACAQPGYPWPLYGNEGLWAGWLLSPYSGRPRWGPTDPRRRSRNLG	120		
Qy	121	KVTDLTTCGADLMGYIPLVGAPLGGAAARALAHGVRVLEGDVNVATGNLPGCSFSIFLLA	180		
Db	121	KVTDLTTCGADLMGYIPLVGAPLGGVARRALAHGVRVLEGDVNVATGNLPGCSFSIFLLA	180		
Qy	181	LLSLCTIPASAYEVRNVSGIYHVNTDCSNSISIVYEAADVMHTPGCVPCVOEGNSSRCWV	240		
Db	181	LLSLCTVPVSTYEVNRNVSGVYHVNTDCSNSISIVYEAADVMHTPGCVPCVREGNSSRCWV	240		
Qy	241	ALPTTLAARNASVPTTIRRHVDLLVGTAAAFCSAMYVDGLCGSIFLVSQLFTSPRRHET	300		
Db	241	ALPTTLAARNSSIPPTTIRRHVDLLVGAALACSAMYVDGLCGSVFLVSQVLTFTSPRRYET	300		
Qy	301	VQDNCISYIPGHVSGHRMAMDMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLGAGLA	360		
Db	301	VQDNCISYIPGHVSGHRMAMDMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLGAGLA	360		
Qy	361	YYSMVGNWAKVLIIVALLFAGVDGTHTRVAGHTTSGTSLFSSGASQKIQLVNTNGSW	420		
Db	361	YYSMVGNWAKVLIIVMLLFAGVDGTTVSGGHASQITRGVTSFFSPGSAQKIQLVNTNGSW	420		
Qy	421	HINRTALNCNDSLOTGFFAALFYAHKFNSSCCPERMASCRIPIIDWFAQGWGPIYTKPNSS	480		
Db	421	HINRTALNCNESINTGFFAALFYANKFNSSCCSERMASCRIPIIDWFAQGWGPIYTHAESRSS	480		
Qy	481	DQPYCHYAPRPGVVPASQVCGPVYCTFSPVVGTTDRSGVPTYSWGENETDVLIN	540		
Db	481	DQPYCHYAPQCGIVPALQVCGPVYCTFSPVVGTTDRFGAFTYNNGANETDVLIN	540		
Qy	541	NTRPQGNWFCGTWNSTGFTKTCGGPPCNIGGVNRTLICTPTDCFRKHPEATYTKCGSG	600		
Db	541	NTRPQGTWFCGTWNSTGFTKTCGGPPCNIGGVGNKTLCTPTDCFRKHPEATYTKCGSG	600		
Qy	601	PWLTPRCIVDYPYRLWHYPCTINFSIFKVRMYVGVGVEHRLNAACNWTGERCNLEDRDS	660		
Db	601	PWLTPRCVMHYPYRLWHYPCTVNFETIFKIRMYVGVGVEHRLNAACNWTGERCNLEDRDS	660		
Qy	661	ELSPILLSTTEWQILPCAFPTTLPALSTGLIHLHQNIVDVQVLYGVGSFAIKWEYIL	720		
Db	661	ELSPILLSTTEWQILPCFTTLPALSTGLIHLHQNIVDVQVLYGVGSFAIKWEYIL	720		

Qy	721	LLFLLIADARVCA	CLWMMLIIAQEA	ALENLVNLNAASVAGAHGILSFLVFFCAAWYIKG	780
Db	721	LLFLLIADARVCA	CLWMMLIIAQEA	ALENLVNLNAASVAGAHGVLFFLVFFCAAWYIKG	780
Qy	781	RL	782		
Db	781	RL	782		
RESULT 15					
S19875					
genome polypeptin - hepatitis C virus (isolate JK3) (fragment)					
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein					
C;Species: hepatitis C virus					
A;Variety: isolate JK3					
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004					
C;Accession: S19875					
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.					
submitted to the EMBL Data Library, September 1991					
A;Description: Sequence analysis of putative structural regions of Hepatitis C Virus seq					
A;Reference number: S18029					
A;Accession: S19875					
A;Molecule type: genomic RNA					
A;Residues: 1-782 <HON>					
A;Cross-references: UNIPROT:Q68951; UNIPARC:UPI00000F03E9; EMBL:X61592; NID:g59482; PIDD					
A;Experimental source: isolate JK3					
C;Superfamily: hepatitis C virus genome polyprotein					
C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural					
F;1-191/Product: core protein #status predicted <MAT1>					
F;192-383/Product: envelope protein 1 #status predicted <MAT2>					
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>					
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>					
Query Match 24.5%; Score 3922; DB 2; Length 782;					
Best Local Similarity 89.8%; Pred. No. 9.9e-230;					
Matches 702; Conservative 37; Mismatches 43; Indels 0; Gaps 0;					
Qy	1	MSNPKPQRTKNTNRR	PODVKPPGGQIVGGVYLLPRRGPRGLGVRA	TKASERSQPRG	60
Db	1	MSNPKPQRTKNTNRR	PODVKPPGGQIVGGVYLLPRRGPRGLGVRA	TKASERSQPRG	60
Qy	61	RRQIPKARPEGRAWAQ	CPWPLVYNEGLGWAGWLLSPRGSRPSG	PGTDPRRSRNLG	120
Db	61	RRQIPKARPEGRAWAQ	CPWPLVYNEGLGWAGWLLSPRGSRPSG	PGTDPRRSRNLG	120
Qy	121	KVTDLTTCGFADLMGY	ILPLVGA PLGGGAARALAHGVRVLEDGVN	YATGNLPGCSFSIFLLA	180
Db	121	KVTDLTTCGFADLMGY	ILPLVGA PLGGGAARALAHGVRVLEDGVN	YATGNLPGCSFSIFLLA	180
Qy	181	LLSCLTIIPASAYEVR	NVSGIYHYVTNDCSNSIIVYEAADVMHT	PGVPCVQEGNSRCWV	240
Db	181	LLSCLTIIPASAYEVR	NVSGIYHYVTNDCSNSIIVYEAADVMHT	PGVPCVQEGNSRCWV	240
Qy	241	ALPTTLAARNASVPT	TTTIRRHVDLLVGTAAFCASAMVVDLCGS	FELVSQLTFSRRHET	300
Db	241	ALPTTLAARNASVPT	TTTIRRHVDLLVGTAAFCASAMVVDLCGS	FELVSQLTFSRRHET	300
Qy	301	VQDCNCSIYPGHVSG	HRMADMMNWSPTTALVVSQLLRIPQAVDM	VAGAHGVLVAGLA	360
Db	301	VQDCNCSIYPGHVSG	HRMADMMNWSPTTALVVSQLLRIPQAVDM	VAGAHGVLVAGLA	360
Qy	361	YYSMVGNAKVLIV	ALLFAGVDGETHTTGRVAGHTTSGFTSLF	SSGASQIKQLVNTGNSW	420
Db	361	YYSMVGNAKVLIV	ALLFAGVDGETHTTGRVAGHTTSGFTSLF	SSGASQIKQLVNTGNSW	420
Qy	421	HINRTALNCNDSI	QTFGFAALFYAHKFNSSGCCPERMASCR	PIDWFAOGWGPIVTKPNSS	480
Db	421	HINRTALNCNDSI	QTFGFAALFYAHKFNSSGCCPERMASCR	PIDWFAOGWGPIVTKPNSS	480
Qy	481	DQRPYCHWYAPRC	GVVPASQVCGPVYCFTPSPVVGTDRSG	VPTYSWGENETDVMLN	540
Db	481	DQRPYCHWYAPRC	GVVPASQVCGPVYCFTPSPVVGTDRSG	VPTYSWGENETDVMLN	540

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 08:16:22 ; Search time 314 Seconds
(without alignments)
4211.879 Million cell updates/sec

Title: US-09-662-454-3
Perfect score: 16009
Sequence: 1 MSTNPKPQKTKRNTNRQ.....FPLCLLLSVGVGYLLPNR 3010

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1990s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16009	100.0	3010	2 AAW98022	Aaw98022 Infectiou
2	16009	100.0	3010	4 AAB31170	Aab31170 Amino aci
3	16009	100.0	3010	8 ADO36227	Ado36227 Hepatitis
4	16009	100.0	3010	8 ADO79401	Ado79401 Hepatitis
5	15963	99.7	3010	4 AAB59174	Aab59174 Protein e
6	15772	98.5	3010	9 ADX40786	Adx40786 HCV polym
7	15367	96.0	3010	7 ADF88597	Adf88597 Hepatitis
8	15367	96.0	3010	9 ADX40808	Adx40808 HCV polym
9	15358	95.9	3010	2 AAR68864	Aar68864 Hepatitis
10	15355	95.9	3010	2 AAR68622	Aar68622 HCV prote
11	15332	95.8	3010	2 AAR68622	Aar68622 HCV prote
12	15323	95.7	3010	5 ABG32451	Abg32451 Hepatitis
13	15319	95.7	3010	5 ABG32460	Abg32460 Hepatitis
14	15319	95.7	3010	5 ABG32454	Abg32454 Hepatitis
15	15318	95.7	3010	5 ABG32461	Abg32461 Hepatitis
16	15317	95.7	3010	5 ABG32458	Abg32458 Hepatitis
17	15317	95.7	3010	5 ABG32459	Abg32459 Hepatitis
18	15317	95.7	3010	5 ABG32455	Abg32455 Hepatitis
19	15317	95.7	3010	5 ABG32457	Abg32457 Hepatitis
20	15317	95.7	3010	5 ABG32452	Abg32452 Hepatitis
21	15316	95.7	3010	5 ABG32453	Abg32453 Hepatitis
22	15312.5	95.6	3011	5 ABG32456	Abg32456 Hepatitis
23	15273	95.4	3010	9 ADX40789	Adx40789 HCV polym
24	15265	95.4	3010	9 ADX40781	Adx40781 HCV polym

25	15253	95.3	3010	2 AAR20111	Aar20111 Non-A, no
26	15253	95.3	3010	2 AAR20091	Aar20091 Non-A, no
27	15253	95.3	3010	9 ADX40806	Adx40806 HCV polym
28	15251	95.3	3010	2 AAY06423	Aay06423 Non-A, no
29	15248	95.2	3010	9 ADX40805	Adx40805 HCV polym
30	15239	95.2	3010	9 ADX40791	Adx40791 HCV polym
31	15236	95.2	3010	9 ADX40811	Adx40811 HCV polym
32	15225	95.1	3010	9 ADX40794	Adx40794 HCV polym
33	15214	95.0	3010	9 ADX40807	Adx40807 HCV polym
34	15197	94.9	3010	9 ADX40818	Adx40818 HCV polym
35	15193	94.9	3010	9 ADX40812	Adx40812 HCV polym
36	15172	94.8	3010	9 ADX40800	Adx40800 HCV polym
37	15142.5	94.6	3091	9 AEB17115	Aeb17115 Hepatitis
38	15126	94.5	3010	9 ADX40816	Adx40816 HCV polym
39	15119	94.4	3010	9 ADX40790	Adx40790 HCV polym
40	15104	94.3	3010	9 ADX40779	Adx40779 HCV polym
41	15075	94.2	3010	9 ADX40793	Adx40793 HCV polym
42	15072	94.1	3010	9 ADX40817	Adx40817 HCV polym
43	15068	94.1	3010	9 ADX40801	Adx40801 HCV polym
44	15067	94.1	3010	2 AAR30616	Aar30616 Polypatti
45	15067	94.1	3010	9 ADX40783	Adx40783 HCV polym

ALIGNMENTS

RESULT 1
AAW98022
ID AAW98022 standard; protein; 3010 AA.
XX AC AAW98022;
XX AC AAW98022;
DT 21-JUN-1999 (first entry)
XX Infectious hepatitis C virus genotype 1b strain HC-J4 protein.
DE Infectious hepatitis C virus genotype 1b strain HC-J4 protein.
XX HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
KW assay; antiviral; virucide.
XX Hepatitis C virus.
XX WO9904008-A2.
XX 28-JAN-1999.
XX 16-JUL-1998; 98WO-US014688.
XX 18-JUL-1997; 97US-0053062P.
PR 27-JAN-1998; 98US-00014416.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
WPI; 1999-132252/11.
DR N-PSDB; AAX24843.
XX New isolated hepatitis C virus nucleic acids - used to develop products
PT for the diagnosis, prevention and treatment of HCV infections and for
PT developing screening assays.
XX Claim 2; Fig 14G-H; 126pp; English.
CC This protein is encoded by the infectious hepatitis C virus (HCV)
CC genotype 1b strain HC-J4 genome (see AAX24833). HC-J4 was obtained from
CC acute phase plasma of a chimpanzee infected with serum containing HC-
CC J4/91. The infectious nucleic acid sequence can be used to produce
CC chimeric genomes (see AAX24833) consisting of the open reading frames of
CC infectious nucleic acid sequences of other genotypes (including genotypes
CC 1-6) and subtypes (such as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV.
CC The invention also relates to the introduction of mutations or deletions
CC into infectious nucleic acid sequences in order to produce an attenuated
CC HCV virus suitable for vaccine development. Infectious nucleic acid

sequences can also be used to produce attenuated virus via passage in vitro or in vivo of the viruses produced by transfection of a host cell with the infectious nucleic acid sequence. Vaccines comprising one or more polypeptides made from the infectious nucleic acid sequence are used to immunize mammals, especially humans, against hepatitis C. The nucleic acid sequences can also be used to induce protective immunity against the virus. The nucleic acid sequences or their encoded proteases (e.g. NS3 protease) can additionally be used to develop screening assays to identify antiviral agents for HCV

XX Sequence 3010 AA;

Query Match 100.0%; Score 16009; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSNTPKPQKTYRNTNRNRRQDVKFPGGQIVGGVILLPRRGRPLGVRAIRKASERSQPRG	60
DB	1	MSNTPKPQKTYRNTNRNRRQDVKFPGGQIVGGVILLPRRGRPLGVRAIRKASERSQPRG	60
QY	61	RRQPIKARRPRGRAWAQQGYWPPLYGNEGLGWAGWLLSPRGRSPSWGPTDPRRSRNLG	120
DB	61	RRQPIKARRPRGRAWAQQGYWPPLYGNEGLGWAGWLLSPRGRSPSWGPTDPRRSRNLG	120
QY	121	KVIDTLTCGFADLMGYIPLVGAFLGGAARALAHGVRLVDGYNVYATGNLPGCSFSIFLLA	180
DB	121	KVIDTLTCGFADLMGYIPLVGAFLGGAARALAHGVRLVDGYNVYATGNLPGCSFSIFLLA	180
QY	181	LLSCLTIPASAEVRNVSIGYHVTNDCSNSISIVYEAAADVMTPGVCVQEGNSRCRW	240
DB	181	LLSCLTIPASAEVRNVSIGYHVTNDCSNSISIVYEAAADVMTPGVCVQEGNSRCRW	240
QY	241	ALTPTLAARNASVPTTIRRHVDLVGTAAFCSAMVVGDLCSIFLVSLQTFSPRRHET	300
DB	241	ALTPTLAARNASVPTTIRRHVDLVGTAAFCSAMVVGDLCSIFLVSLQTFSPRRHET	300
QY	301	VQDCNCSIPYGHVSHRMAWDMNWSPTTALVVSQILRIPOAVDMVAGAHGWLAGLA	360
DB	301	VQDCNCSIPYGHVSHRMAWDMNWSPTTALVVSQILRIPOAVDMVAGAHGWLAGLA	360
QY	361	YYSVMGNWAKVILVALLFAGVDETHTTGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW	420
DB	361	YYSVMGNWAKVILVALLFAGVDETHTTGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW	420
QY	421	HNRTALNCNDSLOCGFFAALFYAHKFNSSGCPERMASCRPDWFAQGGPITYTKPNSS	480
DB	421	HNRTALNCNDSLOCGFFAALFYAHKFNSSGCPERMASCRPDWFAQGGPITYTKPNSS	480
QY	481	DORPYCWHYAPRCGVVPASQVCGPVYCTPTSPVVVGTDRSGVPTYSWGENETDVMILN	540
DB	481	DORPYCWHYAPRCGVVPASQVCGPVYCTPTSPVVVGTDRSGVPTYSWGENETDVMILN	540
QY	541	NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGVGNRTLICPTDCPRKHPEATYTKCGSG	600
DB	541	NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGVGNRTLICPTDCPRKHPEATYTKCGSG	600
QY	601	PMLTPRCLVDYPRYLWHYECTLNFSIFKVRMYVGGVEHRLNAAACNWRGERCNLEDRRS	660
DB	601	PMLTPRCLVDYPRYLWHYECTLNFSIFKVRMYVGGVEHRLNAAACNWRGERCNLEDRRS	660
QY	661	ELSPILLSTTEWQILPCAFTTLPALSTGLIHLHQNIVDVQYLVYGVSFAIKWEYIL	720
DB	661	ELSPILLSTTEWQILPCAFTTLPALSTGLIHLHQNIVDVQYLVYGVSFAIKWEYIL	720
QY	721	LLFLLADARVACALWMLLIIAQEAALLENVLNAAASVAGAHGILLSFLVFFCAAWYIKG	780
DB	721	LLFLLADARVACALWMLLIIAQEAALLENVLNAAASVAGAHGILLSFLVFFCAAWYIKG	780
QY	781	RLAPGAAYAFYGVWPLLLLLLLALPPRAYALDREMAASCAGVLVGLVFTLSPYKVKFLT	840
DB	781	RLAPGAAYAFYGVWPLLLLLLLALPPRAYALDREMAASCAGVLVGLVFTLSPYKVKFLT	840
QY	841	RLIWMLOQYFITRAEAHMVWPPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGPLM	900

841	RLIWMLOQYFITRAEAHMVWPPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGPLM	900	
QY	901	VLQAGITRPVFEVRAQGLIRACMLVRKVAGGHYVQVMFKLGALGTGVVYNHLTPLRDWA	960
DB	901	VLQAGITRPVFEVRAQGLIRACMLVRKVAGGHYVQVMFKLGALGTGVVYNHLTPLRDWA	960
QY	961	HAGLRDLAVAVEPVVFSAMETKVIITWGDATACAGDIIILGLPVSARRGKEIFLGPADSLRG	1020
DB	961	HAGLRDLAVAVEPVVFSAMETKVIITWGDATACAGDIIILGLPVSARRGKEIFLGPADSLRG	1020
QY	1021	QGWRLAPITAYSQOTRGVLGCIITSLTGRDKNQVEGEVQVVSTATQSPLATCINGVCWT	1080
DB	1021	QGWRLAPITAYSQOTRGVLGCIITSLTGRDKNQVEGEVQVVSTATQSPLATCINGVCWT	1080
QY	1081	VYHGAGSKTLGPKGPIITQMTYNVDLDLVGWAQPGABSMTPCSCGSSDLYLVTRHADVI	1140
DB	1081	VYHGAGSKTLGPKGPIITQMTYNVDLDLVGWAQPGABSMTPCSCGSSDLYLVTRHADVI	1140
QY	1141	PVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES	1200
DB	1141	PVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES	1200
QY	1201	METTWRSPVFTDNSTPPAVPQTFQVAHLHAPTGSOKSTKVPAAYAAQGYKVLVLPNSVAA	1260
DB	1201	METTWRSPVFTDNSTPPAVPQTFQVAHLHAPTGSOKSTKVPAAYAAQGYKVLVLPNSVAA	1260
QY	1261	TILGFCAYNSKAGHIDPNRTIGVRTITTCGSIITYSYGKFLADGGCGGAYDIIICDECHS	1320
DB	1261	TILGFCAYNSKAGHIDPNRTIGVRTITTCGSIITYSYGKFLADGGCGGAYDIIICDECHS	1320
QY	1321	TDSTTILGTVLDOAETAGARLVLATATPGSVTVVPHNIEEIGLSNNGEIPFYGKAI	1380
DB	1321	TDSTTILGTVLDOAETAGARLVLATATPGSVTVVPHNIEEIGLSNNGEIPFYGKAI	1380
QY	1381	PIEAIKGRHLIFCHSKKKDELAALKLTGLNLNAVAYRGLDVSVIPPIGDVVVATDAL	1440
DB	1381	PIEAIKGRHLIFCHSKKKDELAALKLTGLNLNAVAYRGLDVSVIPPIGDVVVATDAL	1440
QY	1441	MTGFTGDFSDVDNCNTCVTVDFSLDPTFTTETTTVPQDAVRSQRGRGTGRSGIYR	1500
DB	1441	MTGFTGDFSDVDNCNTCVTVDFSLDPTFTTETTTVPQDAVRSQRGRGTGRSGIYR	1500
QY	1501	FVTPGERPSGDFSSVLCEDYDAGCAWYELTPAETSRLRAYLNTPLGLPVCODHLEFWES	1560
DB	1501	FVTPGERPSGDFSSVLCEDYDAGCAWYELTPAETSRLRAYLNTPLGLPVCODHLEFWES	1560
QY	1561	VFTGLTHIDAFLSQTQAGDNFPYLVAQATVCARAQAPPPSDQMKCLIRLKPTLHG	1620
DB	1561	VFTGLTHIDAFLSQTQAGDNFPYLVAQATVCARAQAPPPSDQMKCLIRLKPTLHG	1620
QY	1621	PTPLLYRILGAVONEVILTHPITKYIMACMSADLEVVWTSTWLVGGVLAALAAAYCLTTGSV	1680
DB	1621	PTPLLYRILGAVONEVILTHPITKYIMACMSADLEVVWTSTWLVGGVLAALAAAYCLTTGSV	1680
QY	1681	VTVGRIILSGKPAVVDREVLYQBEFDEMEECASQLPYTEQGMQLABQFKQKALGLQATAT	1740
DB	1681	VTVGRIILSGKPAVVDREVLYQBEFDEMEECASQLPYTEQGMQLABQFKQKALGLQATAT	1740
QY	1741	KQAEAAAAPVESKWRALETFWAKHWNFIISGIQYLAGLSTLPCNPAIASIAMAFTASITSP	1800
DB	1741	KQAEAAAAPVESKWRALETFWAKHWNFIISGIQYLAGLSTLPCNPAIASIAMAFTASITSP	1800
QY	1801	LTTQNTLILFNILGGWVAALAPPASAAPVAGIAGAAVGSIGLKVLDIILAGYAGVA	1860
DB	1801	LTTQNTLILFNILGGWVAALAPPASAAPVAGIAGAAVGSIGLKVLDIILAGYAGVA	1860
QY	1861	GALVAFKWSGVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNNRLI	1920
DB	1861	GALVAFKWSGVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNNRLI	1920
QY	1921	AFASRGNVSPHYPPESDAAAARVTQIILSSLTITQLKRLHQNEDCSTPCSGSWLRDV	1980

Db 1921 AFASRGHVSPHYVSPESDAAAVTQILSSLTITQLLKRHLQWINEDECSTPCSGSWLRDV 1980
Qy 1981 WDWICTVLTDFKTLQSLKLPRLPGVPFLSCQRYGKVGWRGDGIMQMTCTPCGAQIAGHVK 2040
Db 1981 WDWICTVLTDFKTLQSLKLPRLPGVPFLSCQRYGKVGWRGDGIMQMTCTPCGAQIAGHVK 2040
Qy 2041 NGSMRIVGPRTCNTWHTGTFPIINAYTTGCTPSPAPNYSALWRVAEEVEVTRVGDHF 2100
Db 2041 NGSMRIVGPRTCNTWHTGTFPIINAYTTGCTPSPAPNYSALWRVAEEVEVTRVGDHF 2100
Qy 2101 YVTGMTTDNVKCPQVPAPEFFTEVDGVRHLRRYAPACKPLLRDVTFOVGLNVLVGSOL 2160
Db 2101 YVTGMTTDNVKCPQVPAPEFFTEVDGVRHLRRYAPACKPLLRDVTFOVGLNVLVGSOL 2160
Qy 2161 PCPEPDVTVLTSMLTDPSSHITAEAKRLRARGSPPSLASSASQSAPSLKATCTTHHD 2220
Db 2161 PCPEPDVTVLTSMLTDPSSHITAEAKRLRARGSPPSLASSASQSAPSLKATCTTHHD 2220
Qy 2221 SPADLIEANLLWRQEMGNI TRVESENKVILDSFEPLHAEGDEREISVAAEILKRSRK 2280
Db 2221 SPADLIEANLLWRQEMGNI TRVESENKVILDSFEPLHAEGDEREISVAAEILKRSRK 2280
Qy 2281 FPSALPIWARPDPNPPLESKQDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
Db 2281 FPSALPIWARPDPNPPLESKQDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
Qy 2341 SSALAEIATKTFGSSGSADVSGTATALPOLASDGDGKSDVESYSSMPLEGEFGDPL 2400
Db 2341 SSALAEIATKTFGSSGSADVSGTATALPOLASDGDGKSDVESYSSMPLEGEFGDPL 2400
Qy 2401 SDGSWSTVSEASDVVCCSVYTWGALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
Db 2401 SDGSWSTVSEASDVVCCSVYTWGALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
Qy 2461 SRSASLROKVTFRLOVLDHVDLKEWKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
Db 2461 SRSASLROKVTFRLOVLDHVDLKEWKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
Qy 2521 GYGAKDVRLNSSRAVNHRSWEDLLETETPIDITIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVRLNSSRAVNHRSWEDLLETETPIDITIMAKSEVFCVQPEKGRKPARLIV 2580
Qy 2581 FPDIGVRVCERKALYDVVSTLPQAVMGSSYGFQSPQRVEFLVNTWKSXKCPMGFSYDT 2640
Db 2581 FPDIGVRVCERKALYDVVSTLPQAVMGSSYGFQSPQRVEFLVNTWKSXKCPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNGCYRRCA 2700
Db 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNGCYRRCA 2700
Qy 2701 SGVLTTSCGNTLTCLYKATACRAKLODCTMLVNGDDLVIICISAGTQEDAAALRAFTE 2760
Db 2701 SGVLTTSCGNTLTCLYKATACRAKLODCTMLVNGDDLVIICISAGTQEDAAALRAFTE 2760
Qy 2761 AMTRYSAPPDPPOPEYDLELITSCSSNVSVAHASGRVYVYLTDRPTPLARAWEYAR 2820
Db 2761 AMTRYSAPPDPPOPEYDLELITSCSSNVSVAHASGRVYVYLTDRPTPLARAWEYAR 2820
Qy 2821 HTPINSWLGNIIMVAPTLWARMILMTHFFSILLAQEKLDCQIYGACYSIEPLDLPQ 2880
Db 2821 HTPINSWLGNIIMVAPTLWARMILMTHFFSILLAQEKLDCQIYGACYSIEPLDLPQ 2880
Qy 2881 IIERLHGLSATLHYSYSGEINRVASCLRLGVPLRTWRHARSVRAKLISQGGRAATC 2940
Db 2881 IIERLHGLSATLHYSYSGEINRVASCLRLGVPLRTWRHARSVRAKLISQGGRAATC 2940
Qy 2941 GRYPFNWAVRKLKLTPIPAASQDLSCGFVAGYSGGDIYHSLSRARPRWPLCLLLLSV 3000
Db 2941 GRYPFNWAVRKLKLTPIPAASQDLSCGFVAGYSGGDIYHSLSRARPRWPLCLLLLSV 3000
Qy 3001 GVGIVLLPNR 3010
Db 3001 GVGIVLLPNR 3010

RESULT 2

AAB31170
ID AAB31170 standard; protein; 3010 AA.

XX AAB31170;
XX AC
XX 02-APR-2001 (first entry)
XX DT

XX Amino acid sequence of a hepatitis C virus (HCV) clone genotype 1b.

DE Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
XX HCV; vaccine; viral inhibitor; antiviral.

XX Hepatitis C virus.

XX WO200075352-A2.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015527.

XX 04-JUN-1999; 99US-0137817P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Nam J, Bukh J, Emerson SU, Purcell RH;

XX WPI; 2001-071081/08.

XX N-PSDB; AAC86939.

XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome
PT in which the (non-)structural region has been replaced by hepatitis C
PT virus (HCV) genome useful for treating or preventing HCV signs and
PT symptoms.

XX Disclosure; Fig 4G-H; 97pp; English.

XX The specification describes a nucleic acid comprising a chimeric virus
CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
CC the (non-)structural region has been replaced by the (non-)structural
CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC the chimeric virus and the chimeric virus are useful for identifying cell
CC lines capable of supporting the replication of these chimeric viruses, in
CC screening for neutralizing antibodies to HCV of different genotypes, in
CC the production of HCV-BVDV virions, for the development of inactivated or
CC attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
CC molecular properties of HCV indirectly in vitro, and in identifying
CC inhibitors of viral enzyme activity which would be useful as antiviral
CC agents. Formulations or compositions comprising the chimeric viruses may
CC be used to treat or prevent the signs and symptoms of HCV. The present
CC sequence is encoded by a HCV clone, which is used to construct chimeric
CC nucleic acids of the invention

XX Sequence 3010 AA;

Query Match 100.0%; Score 16009; DB 4; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKPPGGQIVGGVYLLPRRGRPLGVRATRKASERSQPRG 60

Db 1 MSTNPKPQKTKRNTNRRPQDVKPPGGQIVGGVYLLPRRGRPLGVRATRKASERSQPRG 60

Qy 61 RRQPIPKARREPEGRAWQPGYPWPLYGNEGLGWAGWLLSPRGSRPSPWGPTDPPRRSRNLG 120

Db 61 RRQPIPKARREPEGRAWQPGYPWPLYGNEGLGWAGWLLSPRGSRPSPWGPTDPPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVCAPIGGARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPLVCAPIGGARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180

Db 2341 SSALAEATKTFSGSSAVDSGTATAPDLASDDGKSDVESYSSMPLEGEPPDL 2400
Qy 2401 SDGWSVSEASBDDVCCSNYSYTWTCALITPCAAEESKLPINPLSNLLRHHNMVYATT 2460
Db 2401 SDGWSVSEASBDDVCCSNYSYTWTCALITPCAAEESKLPINPLSNLLRHHNMVYATT 2460
Qy 2461 SRSASLRQKKVTFORLOVDHEDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
Db 2461 SRSASLRQKKVTFORLOVDHEDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
Qy 2521 GYGAKDVNRNLSRAVNHRSWEDLLEDTPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVNRNLSRAVNHRSWEDLLEDTPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Qy 2581 FPDIGVRCERKALYDVVSTLPQAVMGSSYGFQSPQRVEFLVNTWKSCKPMGFSDYT 2640
Db 2581 FPDIGVRCERKALYDVVSTLPQAVMGSSYGFQSPQRVEFLVNTWKSCKPMGFSDYT 2640
Qy 2641 RCFDSTVTESDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCA 2700
Db 2641 RCFDSTVTESDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCA 2700
Qy 2701 SGVLTTSCGNLTCTYLKATACRAKLODCTMLVNGDDLVIICESAGTOEDAAALRAFT 2760
Db 2701 SGVLTTSCGNLTCTYLKATACRAKLODCTMLVNGDDLVIICESAGTOEDAAALRAFT 2760
Qy 2761 AMTRYSPAGPPPOPEVDLELITSCSNVSVAHDSAGKRVVYLTRDPTTPLARAWEATAR 2820
Db 2761 AMTRYSPAGPPPOPEVDLELITSCSNVSVAHDSAGKRVVYLTRDPTTPLARAWEATAR 2820
Qy 2821 HTPINSLGNIIMYAPTARMILMTWFFSILLAOEQLKALDCQIYGACYSIPLDLQ 2880
Db 2821 HTPINSLGNIIMYAPTARMILMTWFFSILLAOEQLKALDCQIYGACYSIPLDLQ 2880
Qy 2881 IIERLHGLSAPTLHSYSPGEINRVASCLRLGVPPPLTRWHRARSVRAKLLSQGGRATC 2940
Db 2881 IIERLHGLSAPTLHSYSPGEINRVASCLRLGVPPPLTRWHRARSVRAKLLSQGGRATC 2940
Qy 2941 GRYLFPNAVTRTKLKTPIPAASQLDLSGWFFVAGYSGGDIYHLSRARPRWFLCLLLSV 3000
Db 2941 GRYLFPNAVTRTKLKTPIPAASQLDLSGWFFVAGYSGGDIYHLSRARPRWFLCLLLSV 3000
Qy 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010

RESULT 3
ADO36227

ID ADO36227 standard; protein; 3010 AA.

XX AC ADO36227;

XX DT 26-AUG-2004 (first entry)

XX DE Hepatitis C virus (HCV) J4L6 wild-type polyprotein.

XX KW hepatotropic; virucide; vaccine; gene therapy; vaccine;
KW Hepatitis C virus; HCV; core protein; HCV infection; vaccination;
KW polyprotein.

XX OS Hepatitis C virus.

XX PN WO2004046175-A1.

XX PD 03-JUN-2004.

XX PF 13-NOV-2003; 2003WO-EP012793.

XX PR 15-NOV-2002; 2002GB-00026722.

XX PA (GLAX) GLAXO GROUP LTD.

XX

Pl. Brett S, Hamblin PA, Ogilvie L;

XX WPI; 2004-420613/39.

XX N-PSDB; ADO36222.

PT New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
the polypeptide sequences of the HCV core and at least one other HCV
protein, for use in medicine, particularly for manufacturing a medicament
for treating HCV.

XX Disclosure; Fig 6; 78pp; English.

XX The invention describes a polynucleotide vaccine comprising a
polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
protein and at least 1 other HCV protein, and causes expression of the
proteins in cells (in which (S1) has been mutated or positioned relative
to the polynucleotide sequence encoding the other HCV protein, so that
the negative effect of the Core protein on expression of the other HCV
protein is reduced). Also described are: a method of preventing or
treating an HCV infection in a mammal, comprising administering the
vaccine cited above to a mammal; and a method of vaccination of an
individual, comprising taking a polynucleotide vaccine as cited above,
coating the polynucleotide onto gold beads and delivering the gold beads
into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
antibodies used in the methods, are also disclosed. The polynucleotide
vaccine is useful in the manufacture of a medicament for the treatment of
HCV. This is the amino acid sequence of the wild type HCV polyprotein.

XX Sequence 3010 AA;

Query Match 100.0%; Score 16009; DB 8; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRGRLGVTRATRKASERSQPRG 60

Db 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRGRLGVTRATRKASERSQPRG 60

Qy 61 RROPIPKARPEGRANAQPGYPWPLVNGEGLWAGWLLSPRGRSPSGPTDPRRRSNLG 120

Db 61 RROPIPKARPEGRANAQPGYPWPLVNGEGLWAGWLLSPRGRSPSGPTDPRRRSNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGA PLGGAARALAHGVRVLEDDGVNATGNLPGCSFSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPLVGA PLGGAARALAHGVRVLEDDGVNATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTIPASAYEVRNVSGIYHVNTDCSNSSIVYEADVIHMTPGCVQEGNSSRCWV 240

Db 181 LLSCLTIPASAYEVRNVSGIYHVNTDCSNSSIVYEADVIHMTPGCVQEGNSSRCWV 240

Qy 241 ALPTLAARNASVPTTIRRHVDLLVGTAAFCSNMYVGDLCGSIFLYSQLFTFSPRHET 300

Db 241 ALPTLAARNASVPTTIRRHVDLLVGTAAFCSNMYVGDLCGSIFLYSQLFTFSPRHET 300

Qy 301 VQDCNCSIYPGHVSGHRMAMWMMNMSPTTALVVSQLLRIPQAVDMVAGAHVGLAGLA 360

Db 301 VQDCNCSIYPGHVSGHRMAMWMMNMSPTTALVVSQLLRIPQAVDMVAGAHVGLAGLA 360

Qy 361 YYSVGNWNAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW 420

Db 361 YYSVGNWNAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW 420

Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCGPERMASCRIPIWFAQGWGPITYTKPNSS 480

Db 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCGPERMASCRIPIWFAQGWGPITYTKPNSS 480

Qy 481 DQRPYCHYAPRCGVVPASQVCGPVYCFPTSPVWVGTTRDSGVPTTYSWGENETDVMNLN 540

Db 481 DQRPYCHYAPRCGVVPASQVCGPVYCFPTSPVWVGTTRDSGVPTTYSWGENETDVMNLN 540

Qy 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLCPTDCFRKRPATYTKCGSG 600

Db 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLCPTDCFRKRPATYTKCGSG 600

Qy	601	PWLTPRCLVDYPYRLMHPYCTLNFSIFKVRMYVGGVEHRLNAAACNWTGRGCNLEDRDRS	660
Db	601	PWLTPRCLVDYPYRLMHPYCTLNFSIFKVRMYVGGVEHRLNAAACNWTGRGCNLEDRDRS	660
Qy	661	ELSPLLSTTTEWQILPCAFPTTLPALSTGLIHLHQNIVDVOYLKVGSGAFVSPFAIKWEYIL	720
Db	661	ELSPLLSTTTEWQILPCAFPTTLPALSTGLIHLHQNIVDVOYLKVGSGAFVSPFAIKWEYIL	720
Qy	721	LLFLLLDADARVCACLWMMLLIAQAEAALENLVINAASVAGAGHILSLFVFFCAANVIKG	780
Db	721	LLFLLLDADARVCACLWMMLLIAQAEAALENLVINAASVAGAGHILSLFVFFCAANVIKG	780
Qy	781	RLAPGAAYAFYGVWPLLLLLLALPPRAYALDREMAASC GGAVLVLGLVFLTSPYKVFLT	840
Db	781	RLAPGAAYAFYGVWPLLLLLLALPPRAYALDREMAASC GGAVLVLGLVFLTSPYKVFLT	840
Qy	841	RLIWLQYFITRAEAAHQVWVPPUNVRGGDAIILLTCAVHPELIIDITIKLLAILGLPLM	900
Db	841	RLIWLQYFITRAEAAHQVWVPPUNVRGGDAIILLTCAVHPELIIDITIKLLAILGLPLM	900
Qy	901	VLQAGITRVYFVRAQGLIRACMLVRKVAGGHYQWVFMKLGALTGYVYNHLTPLRDWA	960
Db	901	VLQAGITRVYFVRAQGLIRACMLVRKVAGGHYQWVFMKLGALTGYVYNHLTPLRDWA	960
Qy	961	HAGURDLA VEPVVFSAEMTKVITWGADTAAACGDIILGLPVSAARRGKEIFLGPADSLG	1020
Db	961	HAGURDLA VEPVVFSAEMTKVITWGADTAAACGDIILGLPVSAARRGKEIFLGPADSLG	1020
Qy	1021	QGWRLLAPI TAYSQOTRGVLGCIITSLTGRDKQVGEQVQVSTATQSFATCINGVCWT	1080
Db	1021	QGWRLLAPI TAYSQOTRGVLGCIITSLTGRDKQVGEQVQVSTATQSFATCINGVCWT	1080
Qy	1081	VYHGAGSKTLGPKGPITOMYTNVDLDLVGWOAPPGARSMTPCSCGSSDLYLVTRHADVI	1140
Db	1081	VYHGAGSKTLGPKGPITOMYTNVDLDLVGWOAPPGARSMTPCSCGSSDLYLVTRHADVI	1140
Qy	1141	PVRRRGDSRGLLSPRPVSYLKGSSGGPLLCPSGHVGVGFRAAVCTRGVAKAVDFIPVES	1200
Db	1141	PVRRRGDSRGLLSPRPVSYLKGSSGGPLLCPSGHVGVGFRAAVCTRGVAKAVDFIPVES	1200
Qy	1201	METTMRSPVFTDNTSPPAVPQTFQVAHLHAPTTGSGKSTKVPAAAYAAQYKVLVLPNSVAA	1260
Db	1201	METTMRSPVFTDNTSPPAVPQTFQVAHLHAPTTGSGKSTKVPAAAYAAQYKVLVLPNSVAA	1260
Qy	1261	TLGFGAYMSKAHGDIPNIRTVRTITTGSGITSTYTKFLADGCSGGAYDIIICBECHS	1320
Db	1261	TLGFGAYMSKAHGDIPNIRTVRTITTGSGITSTYTKFLADGCSGGAYDIIICBECHS	1320
Qy	1321	TDSITLIGI GVLDOAETAGARLVLATATPPGSVTVPHNIEBIGLSNNGEIPFYKAI	1380
Db	1321	TDSITLIGI GVLDOAETAGARLVLATATPPGSVTVPHNIEBIGLSNNGEIPFYKAI	1380
Qy	1381	PIEAIKGRHILFCHSKKKCDLAAKLTGLGLNAVAYYRGLDVSVIPIIDGVVVVATDAL	1440
Db	1381	PIEAIKGRHILFCHSKKKCDLAAKLTGLGLNAVAYYRGLDVSVIPIIDGVVVVATDAL	1440
Qy	1441	MTGFTGDFDSVIDCNTCTVTQTVDFSLDPTFTIETTTVPQDAVSRQRGRGSGIYR	1500
Db	1441	MTGFTGDFDSVIDCNTCTVTQTVDFSLDPTFTIETTTVPQDAVSRQRGRGSGIYR	1500
Qy	1501	FVTGERPSGMSDFSILCECYDAGCANWELTPAETSURLRAYLNTGRLPCVODHLEFWS	1560
Db	1501	FVTGERPSGMSDFSILCECYDAGCANWELTPAETSURLRAYLNTGRLPCVODHLEFWS	1560
Qy	1561	VFTGLTHIDAHFLSQTQKAGNFPYLVAYQATVCARAQAPPPSMDQMKCLIRLKP TLHG	1620
Db	1561	VFTGLTHIDAHFLSQTQKAGNFPYLVAYQATVCARAQAPPPSMDQMKCLIRLKP TLHG	1620
Qy	1621	PTPLLRYLGAVONEVILTHPTIKYIMACMSADLEVTSTWVLVGGVLAALAAAYCLTTGVS	1680
Db	1621	PTPLLRYLGAVONEVILTHPTIKYIMACMSADLEVTSTWVLVGGVLAALAAAYCLTTGVS	1680

Qy	1681	VI	GR	I	I	S	K	P	A	V	D	P	R	E	V	L	Y	Q	E	F	D	E	M	E	B	E	C	A	S	Q	L	P	I	E	I	C	M	Q	L	A	E	F	K	O	K	A	L	G	L	L	O	T	A	T	1740			
Db	1681	VI	GR	I	I	S	K	P	A	V	D	P	R	E	V	L	Y	Q	E	F	D	E	M	E	B	E	C	A	S	Q	L	P	I	E	I	C	M	Q	L	A	E	F	K	O	K	A	L	G	L	L	O	T	A	T	1740			
Qy	1741	K	O	E	A	A	P	V	V	E	S	K	W	R	A	L	E	T	F	W	A	K	H	M	N	F	I	S	G	I	O	L	A	G	L	S	T	L	P	C	N	P	A	T	A	S	L	T	S	P	1800							
Db	1741	K	O	E	A	A	P	V	V	E	S	K	W	R	A	L	E	T	F	W	A	K	H	M	N	F	I	S	G	I	O	L	A	G	L	S	T	L	P	C	N	P	A	T	A	S	L	T	S	P	1800							
Qy	1801	L	T	T	Q	N	T	L	L	F	N	I	L	G	W	A	A	Q	L	P	S	A	S	A	F	V	G	A	G	I	A	G	A	A	V	G	S	I	G	L	K	V	L	D	I	L	A	G	Y	G	A	V	A	1860				
Db	1801	L	T	T	Q	N	T	L	L	F	N	I	L	G	W	A	A	Q	L	P	S	A	S	A	F	V	G	A	G	I	A	G	A	A	V	G	S	I	G	L	K	V	L	D	I	L	A	G	Y	G	A	V	A	1860				
Qy	1861	G	A	L	V	A	F	K	V	M	S	G	E	V	P	S	T	E	D	L	V	N	L	L	P	A	I	L	S	P	A	L	S	P	A	L	S	P	A	L	S	P	A	L	S	P	A	L	S	P	1920							
Db	1861	G	A	L	V	A	F	K	V	M	S	G	E	V	P	S	T	E	D	L	V	N	L	L	P	A	I	L	S	P	A	L	S	P	A	L	S	P	A	L	S	P	A	L	S	P	A	L	S	P	1920							
Qy	1921	A	F	A	S	R	G	N	H	S	P	T	H	Y	V	P	E	S	D	A	A	A	V	T	O	L	S	L	I	T	T	O	L	L	K	R	L	H	O	M	I	N	E	D	C	S	T	P	C	S	G	S	W	L	R	D	1980	
Db	1921	A	F	A	S	R	G	N	H	S	P	T	H	Y	V	P	E	S	D	A	A	A	V	T	O	L	S	L	I	T	T	O	L	L	K	R	L	H	O	M	I	N	E	D	C	S	T	P	C	S	G	S	W	L	R	D	1980	
Qy	1981	W	D	M	I	C	T	V	L	T	D	F	K	T	W	L	O	S	K	L	P	R	L	P	G	V	P	F	L	S	C	O	R	G	Y	K	G	W	R	G	D	I	M	O	T	T	C	P	C	Q	A	I	A	G	H	V	K	2040
Db	1981	W	D	M	I	C	T	V	L	T	D	F	K	T	W	L	O	S	K	L	P	R	L	P	G	V	P	F	L	S	C	O	R	G	Y	K	G	W	R	G	D	I	M	O	T	T	C	P	C	Q	A	I	A	G	H	V	K	2040
Qy	2041	N	G	S	M	R	I	V	G	P	R	T	C	N	T	W	H	G	T	F	P	I	N	A	T	T	G	P	C	T	P	S	P	A	N	S	R	A	L	M	R	V	A	E	E	Y	V	E	T	R	V	G	D	F	H	2100		
Db	2041	N	G	S	M	R	I	V	G	P	R	T	C	N	T	W	H	G	T	F	P	I	N	A	T	T	G	P	C	T	P	S	P	A	N	S	R	A	L	M	R	V	A	E	E	Y	V	E	T	R	V	G						

Db 2761 AMTRYSAAPGDPPEYDLELITSCSSNVSAHDASGKVVYLTDRPTTPLARAAMETAR 2820
Qy 2821 HTPINSWLGNIIMYAPTILWARMILMTHPESIIILAOEKLKALDCOILYGACYIEPLDLPO 2880
Db 2821 HTPINSWLGNIIMYAPTILWARMILMTHPESIIILAOEKLKALDCOILYGACYIEPLDLPO 2880
Qy 2881 IIERHLGSAFTLHYSYSGEINRVASCLRKLGVPPLRTWRHRARSVRKLLSQGGRAATC 2940
Db 2881 IIERHLGSAFTLHYSYSGEINRVASCLRKLGVPPLRTWRHRARSVRKLLSQGGRAATC 2940
Qy 2941 GRYLFNWAVRTKLTPIPAASOLDLSGWFVAGYSGGDIYHLSLRARPWFPLCLLLLSV 3000
Db 2941 GRYLFNWAVRTKLTPIPAASOLDLSGWFVAGYSGGDIYHLSLRARPWFPLCLLLLSV 3000
Qy 3001 GVGIVLLPNR 3010
Db 3001 GVGIVLLPNR 3010

RESULT 4

ID ADO79401 standard; protein; 3010 AA.

XX AC ADO79401;

XX DT 26-AUG-2004 (first entry)

DE DE Hepatitis C virus J4L6 genome wild-type polyprotein.

XX HCV; polyprotein; vaccine; DNA immunisation; hepatotropic; virucide.

XX OS Hepatitis C virus.

OS Synthetic.

XX PN WO2004046176-A1.

XX PD 03-JUN-2004.

XX PF 13-NOV-2003; 2003WO-EP012830.

XX PR 15-NOV-2002; 2002GB-00026722.

XX PA (GLAX) GLAXO GROUP LTD.

XX PI Brett S, Hamblin PA, Ogilvie L;

XX WPI; 2004-420614/39.

DR N-PSDB; ADO79396.

XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins, for use in medicine, in particular for manufacturing a medicament for the treatment of HCV.

XX Disclosure; Fig 6; 79pp; English.

XX The present sequence is the hepatitis C virus (HCV) J4L6 genome wild-type polyprotein sequence. HCV DNA vaccines of the invention comprise a polynucleotide that encodes the HCV proteins Core, NS3, NS4B and NS5B, and does not encode the NS4A and/or NS5A proteins. Preferably, the codon usage of the polynucleotide sequence resembles that of highly expressed human genes. The polynucleotides may encode individual proteins or fusion proteins. Preferred fusions include double fusions between NS4B and NS5B and between Core and NS3. The vaccines are useful for the treatment or prevention of an HCV infection.

XX Sequence 3010 AA;

Query Match 100.0%; Score 16009; DB 8; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKPGGGOIYGVYLLPRRGRLGVRAATKASERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKPGGGOIYGVYLLPRRGRLGVRAATKASERSQPRG 60
Qy 61 RROPIPKARRPEGRAWAOPGYFWPLLYGNEGLGACWLLSPRGSRRPSWGPDPDRRRSRNLG 120
Db 61 RROPIPKARRPEGRAWAOPGYFWPLLYGNEGLGACWLLSPRGSRRPSWGPDPDRRRSRNLG 120
Qy 121 KVIDTLTCGFDLDMGYIPLVGAPLGGAAALAHGVVRVLEEDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFDLDMGYIPLVGAPLGGAAALAHGVVRVLEEDGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTIIPASAYEVNRVSGIYHVTNDCNSSIYVEAADVIMHTPGCVPCVQGENSRRCWV 240
Db 181 LLSCLTIIPASAYEVNRVSGIYHVTNDCNSSIYVEAADVIMHTPGCVPCVQGENSRRCWV 240
Qy 241 ALTPTLAARNASVPTTITRRHVDLLVGTAAAFCSAMVVDLCGSI FLVSQLTFTFSRPHET 300
Db 241 ALTPTLAARNASVPTTITRRHVDLLVGTAAAFCSAMVVDLCGSI FLVSQLTFTFSRPHET 300
Qy 301 VQDCNCSIYPGHVSGHRMADMMMNWSPPTALVVSQLLRIPOAVVDMVAGAHGVLGALA 360
Db 301 VQDCNCSIYPGHVSGHRMADMMMNWSPPTALVVSQLLRIPOAVVDMVAGAHGVLGALA 360
Qy 361 YYSVMGNNAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
Db 361 YYSVMGNNAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSGCGPERMASCRIPTDMFAQGGPIITYTKPNSS 480
Db 421 HINRTALNCNDSLOTGFFAALFYAHKFNSGCGPERMASCRIPTDMFAQGGPIITYTKPNSS 480
Qy 481 DQRPYCMHWAPRPGVVPASQVCGPVYCFPTSPVVVGTTRDSGVPYTSNGENETDVMNLN 540
Db 481 DQRPYCMHWAPRPGVVPASQVCGPVYCFPTSPVVVGTTRDSGVPYTSNGENETDVMNLN 540
Qy 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIIGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Db 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIIGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Qy 601 PMLTPRCLVDYPIRLMHPYCTLNFSIFKVRMYVGGVEHRLNAAACNWTGRGRCNLEDRDRS 660
Db 601 PMLTPRCLVDYPIRLMHPYCTLNFSIFKVRMYVGGVEHRLNAAACNWTGRGRCNLEDRDRS 660
Qy 661 ELSPLLLSTTEWQILPCAFTTLPALSTGLIHLHQNIVDVQYLVGVGSFAFPAIKWEYIL 720
Db 661 ELSPLLLSTTEWQILPCAFTTLPALSTGLIHLHQNIVDVQYLVGVGSFAFPAIKWEYIL 720
Qy 721 LFLLLADARVCACILMMLLIAQAEAALENLVNLNAAASVAGAHGILSFLVFFCAAWYIKG 780
Db 721 LFLLLADARVCACILMMLLIAQAEAALENLVNLNAAASVAGAHGILSFLVFFCAAWYIKG 780
Qy 781 RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCAGVALVGLVFLTSLPYKVFILT 840
Db 781 RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCAGVALVGLVFLTSLPYKVFILT 840
Qy 841 RLTIWLOQYFITRAEAMQVWVPPNLNVRGGRDAIILLTCAVHPELIIDITKLLAILGLPLM 900
Db 841 RLTIWLOQYFITRAEAMQVWVPPNLNVRGGRDAIILLTCAVHPELIIDITKLLAILGLPLM 900
Qy 901 VLOAGITRVPYFVRAOGLIRACMLVRKAGGHVQVWFMKLGALTGYVYVNHITPLRDWA 960
Db 901 VLOAGITRVPYFVRAOGLIRACMLVRKAGGHVQVWFMKLGALTGYVYVNHITPLRDWA 960
Qy 961 HAGLRDLAVAVEPWFVSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HAGLRDLAVAVEPWFVSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Qy 1021 QGWRLLAPITAYSQQTGRGVLGCIITSITGRDKNQVEGEVQVWSTATQSLATCINGVCWT 1080
Db 1021 QGWRLLAPITAYSQQTGRGVLGCIITSITGRDKNQVEGEVQVWSTATQSLATCINGVCWT 1080
Qy 1081 VYHAGSKTLAGPKGPIITQMYTNVDLVLGWAQPPGARSMTWPCSCGSDLLVLTTRHADVI 1140

[illegible]

Db	2161	PCSEPEPDVTVLTSM	LTDPGSHITAE	TAKRR	LARGSP	PPSLASS	QS	AS	PS	LKATCT	THHD	2222																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Qy	2221	SPDADLIEANLLWRQ	BWGGNITR	VESENK	VVILDS	FEPL	HA	E	G	D	E	R	E	I	S	V	A	A	E	I	L	K	R	K	2280																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Db	2221	SPDADLIEANLLWRQ	BWGGNITR	VESENK	VVILDS	FEPL	HA	E	G	D	E	R	E	I	S	V	A	A	E	I	L	K	R	K	2280																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Qy	2281	FPSALPIWARP	DYNPPL	LESWKD	PDVVP	VPVH	GC	PLPT	K	A	P	I	P	P	R	K	R	T	V	L	T	E	S	N	V	2340																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
Db	2281	FPSALPIWARP	DYNPPL	LESWKD	PDVVP	VPVH	GC	PLPT	K	A	P	I	P	P	R	K	R	T	V	L	T	E	S	N	V	2340																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
Qy	2341	SSALAE	LATKTF	SGSS	SAVDS	G	T	A	T	A	L	P	D	L	A	S	D	G	D	G	S	D	V	E	S	Y	S	M	P	L	E	G	E	G	D	P	L	2400																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Db	2341	SSALAE	LATKTF	SGSS	SAVDS	G	T	A	T	A	L	P	D	L	A	S	D	G	D	G	S	D	V	E	S	Y	S	M	P	L	E	G	E	G	D	P	L	2400																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Qy	2401	SDGS	WSTV	EE	ASE	D	V	V	C	S	M	S	T	W	T	G	A	L	I	P	C	A	A	E	S	K	L	P	I	N	P	L	S	N	L	L	R	H	N	M	V	A	T	2460																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Db	2401	SDGS	WSTV	EE	ASE	D	V	V	C	S	M	S	T	W	T	G	A	L	I	P	C	A	A	E	S	K	L	P	I	N	P	L	S	N	L	L	R	H	N	M	V	A	T	2460																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Qy	2461	SR	SAS	L	R	Q	K	V	T	F	D	R	L	Q	V	L	D	D	H	R	V	D	L	K	E	M	K	A	S	T	V	K	A	L	L	S	I	E	B	A	C	K	L	2520																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Db	2461	SR	SAS	L	R	Q	K	V	T	F	D	R	L	Q	V	L	D	D	H	R	V	D	L	K	E	M	K	A	S	T	V	K	A	L	L	S	I	E	B	A	C	K	L	2520																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Qy	2521	G	Y	K	A	D	V	R	N	L	S	S	R	A	V	H	I	R	S	W	E	D	L	L	E	T	P	I	D	T	I	M	A	K	S	E	V	C	V	O	P	E	K	G	R	K	P	A	R	L	I	V	2580																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
Db	2521	G	Y	K	A	D	V	R	N	L	S	S	R	A	V	H	I	R	S	W	E	D	L	L	E	T	P	I	D	T	I	M	A	K	S	E	V	C	V	O	P	E	K	G	R	K	P	A	R	L	I	V	2580																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
Qy	2581	F	P	D	L	G	V	R	V	C	K	M	A	L	I	D	V	S	T	L	P	O	A	V	M	G	S	S	Y	G	F	O	Y	S	P	K	O	R	V	E	F	L	N	T	W	K	S	K	C	P	M	G	F	S	Y	D	T	2640																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
Db	2581	F	P	D	L	G	V	R	V	C	K	M	A	L	I	D	V	S	T	L	P	O	A	V	M	G	S	S	Y	G	F	O	Y	S	P	K	O	R	V	E	F	L	N	T	W	K	S	K	C	P	M	G	F	S	Y	D	T	2640																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
Qy	2641	R	C	F	D	S	T	V	T	E	S	D	I	R	V	E	S	I	Y	Q	C	C	D	L	A	P	E	A	R	Q	A	I	R	S	L	T	E	R	L	I	Y	I	G	G	L	T	N	S	K	G	N	C	G	Y	R	R	C	R	A	2700																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Db	2641	R	C	F	D	S	T	V	T	E	S	D	I	R	V	E	S	I	Y	Q	C	C	D	L	A	P	E	A	R	Q	A	I	R	S	L	T	E	R	L	I	Y	I	G	G	L	T	N	S	K	G	N	C	G	Y	R	R	C	R	A	2700																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Qy	2701	S	G	V	L	T	T	S	C	G	N	T	L	T	C	Y	L	K	A	T	A	C	R	A	A	L	Q	D	C	T	M	L	V	N	G	D	L	V	I	C	E	S	A	G	T	O	E	D	A	A	L	R	A	F	T	E	2760																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Db	2701	S	G	V	L	T	T	S	C	G	N	T	L	T	C	Y	L	K	A	T	A	C	R	A	A	L	Q	D	C	T	M	L	V	N	G	D	L	V	I	C	E	S	A	G	T	O	E	D	A	A	L	R	A	F	T	E	2760																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Qy	2761	A	M	T	R	S	A	P	C	D	P	P	O	P	E	V	D	L	E	I	T	S	C	S	N	S	V	A	H	D	A	S	G	K	V	Y	L	T	R	D	P	T	T	P	L	A	R	A	A	E	T	A	R	2820																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Db	2761	A	M	T	R	S	A	P	C	D	P	P	O	P	E	V	D	L	E	I	T	S	C	S	N	S	V	A	H	D	A	S	G	K	V	Y	L	T	R	D	P	T	T	P	L	A	R	A	A	E	T	A	R	2820																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Qy	2821	H	T	P	I	N	S	W	L	G	N	I	I	M	A	P	T	L	W	A	R	M	L	M	T	H	F	S	I	L	L	A	O	E	K	A	L	D	C	Q	I	Y	G	A	C	S	T	E	P	L	D	L	P	Q	2880																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
Db	2821	H	T	P	I	N	S	W	L	G	N	I	I	M	A	P	T	L	W	A	R	M	L	M	T	H	F	S	I	L	L	A	O	E	K	A	L	D	C	Q	I	Y	G	A	C	S	T	E	P	L	D	L	P	Q	2880																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
Qy	2881	I	I	E	R	L	H	G	L	S	A	F	T	L	H	S	Y	S	P	G	E	I	N	R	V	A	S	C	L	R	K	L	G	V	P	P	L	T	R	W	R	A	R	S	V	R	A	K	L	L	S	O	G	G	R	A	A	T	C	2940																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Db	2881	I	I	E	R	L	H	G	L	S	A	F	T	L	H	S	Y	S	P	G	E	I	N	R	V	A	S	C	L	R	K	L	G	V	P	P	L	T	R	W	R	A	R	S	V	R	A	K	L	L	S	O	G	G	R	A	A	T	C	2940																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
Qy	2941	G	R	V	L	F	N	W	A	R	T	K	L	T	P	I	P	A	A	S	O	L	D	L	S	G	M	F	V	A	G	Y	S	G	G	D	I	Y	H	S	L	S	R	A	R	P	R	F	P	L	C	L	L	L	S	V	3000																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Db	2941	G	R	V	L	F	N	W	A	R	T	K	L	T	P	I	P	A	A	S	O	L	D	L	S	G	M	F	V	A	G	Y	S	G	G	D	I	Y	H	S	L	S	R	A	R	P	R	F	P	L	C	L	L	L	S	V	3000																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
Qy	3001	G	V	G	I	Y	L	L	P	N	R	3010																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						

Db	1681	VIVGRILISLGGKPAVVPDREVLVYQBFDEMECASQLPYIEQMGQAEQFKQKALGLIQTAT	1740
Qy	1741	KQAEAAAPVVESEKRALETETWAKEMWNFIISGIIQVLAGLSTLPGNPALASMAFTASITSP	1800
Db	1741	KQAEAAAPVVESEKRALETETWAKEMWNFIISGIIQVLAGLSTLPGNPALASMAFTASITSP	1800
Qy	1801	LTTQNTLLFNILGWVAAQIAAPPASAAFAFVAGAGIAGAAVSGIGLGKVLVDILAGYGAGVA	1860
Db	1801	LTTQNTLLFNILGWVAAQIAAPPASAAFAFVAGAGIAGAAVSGIGLGKVLVDILAGYGAGVA	1860
Qy	1861	GALVAFKVMSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVMNRLI	1920
Db	1861	GALVAFKVMSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVMNRLI	1920
Qy	1921	AFASRGNHVSTHYVPESDAAARVTQILSSLTITQLLKRHLQWINECDSTPCSGSWLRDV	1980
Db	1921	AFASRGNHVSTHYVPESDAAARVTQILSSLTITQLLKRHLQWINECDSTPCSGSWLRDV	1980
Qy	1981	WDMICTVLTDFTKTLQSKLLPRLPGVPFLSCQRCYKGVWRGDGIMQTTCPGQAQIAGHVK	2040
Db	1981	WDMICTVLTDFTKTLQSKLLPRLPGVPFLSCQRCYKGVWRGDGIMQTTCPGQAQIAGHVK	2040
Qy	2041	NGSMRIIVGPRTCNSNTHGTPPINAYTTGPTCPSPAPNYSRALMRVAABEYEVETRVGDFH	2100
Db	2041	NGSMRIIVGPRTCNSNTHGTPPINAYTTGPTCPSPAPNYSRALMRVAABEYEVETRVGDFH	2100
Qy	2101	YVTGMTDNNVKPCQVPAPEFTEVDGVRHLRYAPACKPLLRDVTFOVGLNOYLVGSOL	2160
Db	2101	YVTGMTDNNVKPCQVPAPEFTEVDGVRHLRYAPACKPLLRDVTFOVGLNOYLVGSOL	2160
Qy	2161	PCEPEPDVTLTSMLTDPSSHITAETAKRRLARGSPPSLASSASSQLSAPSLKATCTTHHD	2220
Db	2161	PCEPEPDVTLTSMLTDPSSHITAETAKRRLARGSPPSLASSASSQLSAPSLKATCTTHHD	2220
Qy	2221	SPDADLLEANLLWRQEMGNITRVESENKVVILDSFEPFLHAEGDEREISVAAEILRLKSRK	2280
Db	2221	SPDADLLEANLLWRQEMGNITRVESENKVVILDSFEPFLHAEGDEREISVAAEILRLKSRK	2280
Qy	2281	FPSALPIWARPDYNPPILLESWKDPDYVPVVGCPPLPTKAPPIPPRRKRTVVLTSNV	2340
Db	2281	FPSALPIWARPDYNPPILLESWKDPDYVPVVGCPPLPTKAPPIPPRRKRTVVLTSNV	2340
Qy	2341	SSALAEATKTFGSGSGSAVDSGTATLPDLASDDGDKGSDVESYSNMPLEGPBGDPDL	2400
Db	2341	SSALAEATKTFGSGSGSAVDSGTATLPDLASDDGDKGSDVESYSNMPLEGPBGDPDL	2400
Qy	2401	SDGWSSTVSEBASEDVVCCSMYSYTWGTALITPCAABERSKLPINPLNSLLRHHNVYATT	2460
Db	2401	SDGWSSTVSEBASEDVVCCSMYSYTWGTALITPCAABERSKLPINPLNSLLRHHNVYATT	2460
Qy	2461	SRSASLRQKVTFFORLOVLDHRYDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF	2520
Db	2461	SRSASLRQKVTFFORLOVLDHRYDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF	2520
Qy	2521	GYGAKDVNLSRANVHRSVWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV	2580
Db	2521	GYGAKDVNLSRANVHRSVWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV	2580
Qy	2581	FPDLGVRVCEKVALYDVVSTLTPQAVMGSSYGFQVSPKQRFVLTNWKSKCPMGFSYDT	2640
Db	2581	FPDLGVRVCEKVALYDVVSTLTPQAVMGSSYGFQVSPKQRFVLTNWKSKCPMGFSYDT	2640
Qy	2641	RCFDSVTVESDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGCQGYRCRA	2700
Db	2641	RCFDSVTVESDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGCQGYRCRA	2700
Qy	2701	SGVLTTSCGNLTLCYLKATAACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAALRAFTE	2760
Db	2701	SGVLTTSCGNLTLCYLKATAACRAAKLODCTMLVNGDDLVIWICESAGTQEDAAALRAFTE	2760
Qy	2761	AMTRYSAFPGPPDPEYDLEILITSCSNVSVVAHDASGRVYVLTFRDPTTPIARAWEFAR	2820
Db	2761	AMTRYSAFPGPPDPEYDLEILITSCSNVSVVAHDASGRVYVLTFRDPTTPIARAWEFAR	2820
Qy	2821	HTPINSWLGNIIMYAPTLMWARMILMTHFFSILLAEQLEKALDCCQIYGACYSIEPLDLPO	2880
Db	2821	HTPINSWLGNIIMYAPTLMWARMILMTHFFSILLAEQLEKALDCCQIYGACYSIEPLDLPO	2880
Qy	2881	IIRERLHGLSAPTLHSHVSPGEINRVASCLRGVPLRTWRHRARSVRAKLLSQGGRAATC	2940
Db	2881	IIRERLHGLSAPTLHSHVSPGEINRVASCLRGVPLRTWRHRARSVRAKLLSQGGRAATC	2940
Qy	2941	GRYLFNWAVRTKLKLTPIPAASQOLDLSGMFVAGYSGDDIYHLSLRARPRWFFPLCLLLSV	3000
Db	2941	GRYLFNWAVRTKLKLTPIPAASQOLDLSGMFVAGYSGDDIYHLSLRARPRWFFPLCLLLSV	3000
Qy	3001	GVGIYLLPNR 3010	
Db	3001	GVGIYLLPNR 3010	
RESULT 6			
ADX40786			
ID	ADX40786 standard; protein; 3010 AA.		
AC	ADX40786;		
XX			
DT	21-APR-2005 (first entry)		
XX			
DE	HCV polymerase protein #9.		
XX			
KW	Immune stimulation; polymerase; enzyme.		
XX			
OS	Hepatitis C virus.		
XX			
PN	WO2005012502-A2.		
XX			
PD	10-FEB-2005.		
XX			
PF	29-MAR-2004; 2004WO-US009510.		
XX			
PR	28-MAR-2003; 2003US-0458026P.		
XX			
PA	(EPIM-) EPIMMUNE INC.		
XX			
PI	Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;		
XX			
DR	WPI; 2005-132661/14.		
XX			
PT	Identifying a candidate peptide epitope, which induces a HLA class I CTL response comprises identifying variants of a peptide epitope 8-11 amino acids in length comprising primary anchor residues of the same HLA class I binding motif.		
XX			
PS	Disclosure; Page 388-440; 458pp; English.		
XX			
CC	The invention relates to a method of identifying a candidate peptide epitope which induces an HLA class I CTL response against variants of the peptide epitope, comprising identifying, from a particular antigen of an infectious agent, variants of a peptide epitope comprising primary anchor residues of the same HLA class I binding motif. The method is useful for identifying a candidate peptide epitope, which induces an HLA class I CTL response against variants of the peptide epitope. This sequence represents an HCV polymerase protein used in the scope of the invention.		
XX			
SQ	Sequence 3010 AA;		
Query Match 98.5%; Score 15772; DB 9; Length 3010;			
Best Local Similarity 98.3%; Pred. No. 0;			
Matches 2959; Conservative 19; Mismatches 32; Indels 0; Gaps 0;			
Qy	1	MSTNPKPQRTKENTRRRPODVKPGGGQIVGGVYLLPRGPRIGVTRATRKASERSQPRG	60
Db	1	MSTNPKPQRTKENTRRRPODVKPGGGQIVGGVYLLPRGPRIGVTRATRKATRSQPRG	60
Qy	61	RRQFIPKARPEGRAWAQPGYPMPLYGNEGIGWAGLLSPRSGRPSWGPTDPRRRSNLIG	120

Db	61	WRQPIKARRPEGRAWAQPGYMPWPLVYGNGLWAGWLLSPGRSPRSGWGTDPARRRNLG	120
Qy	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVLEDDGVNAYATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVLEDDGVNAYATGNLPGCSFSIFLLA	180
Qy	181	LLSCLTIPASAYVRNVSGIYHVNTDCSNSSIVYEADVMHTPGVPCVBOGNSRCWV	240
Db	181	LLSCLTIPASAYVRNVSGIYHVNTDCSNSSIVYEADVMHTPGVPCVBOGNSRCWV	240
Qy	241	ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLCSGIFLVLSOLFSPRRHET	300
Db	241	ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLCSGIFLVLSOLFSPRRHET	300
Qy	301	VQDCNCSIYPGHLSGHRMAMMMWNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGLA	360
Db	301	VQDCNCSIYPGHLSGHRMAMMMWNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGLA	360
Qy	361	YYSWGNWAKVLIVALLFAGVDGETHTRVAGHTTSGFTSLFSSGASQIKLIVNTNGSW	420
Db	361	YYSWGNWAKVLIVALLFAGVDGETHTRVAGHTTSGFTSLFSSGASQIKLIVNTNGSW	420
Qy	421	HINRTALNCNDSIQTFEFAALFYAHKFNSSGCCPERMASCPIDWFAQGWGPITYTKPNS	480
Db	421	HINRTALNCNDSIQTFEFAALFYAHKFNSSGCCPERMASCPIDWFAQGWGPITYTKPNS	480
Qy	481	DQRPYCHYAPRCGVVPASQVCPVYCFPTSPVVVGTTRDSGVPTYSWGENETDVMNLN	540
Db	481	DQRPYCHYAPRCGVVPASQVCPVYCFPTSPVVVGTTRDSGVPTYSWGENETDVMNLN	540
Qy	541	NTRPPQGNWFGCTWMNSTGFTKTCGPPNCNIGVGNRTLICPTDCPRKHPEATYKCGSG	600
Db	541	NTRPPQGNWFGCTWMNSTGFTKTCGPPNCNIGVGNRTLICPTDCPRKHPEATYKCGSG	600
Qy	601	PWLTPRCLVDYPRVRLMHPCTLNPSIFKVRMYGVGVEHRLNAACNWTGERCNLEDRDS	660
Db	601	PWLTPRCLVDYPRVRLMHPCTLNPSIFKVRMYGVGVEHRLNAACNWTGERCNLEDRDS	660
Qy	661	ELSPLLSTTEWQILPCAFTLTPALSTGLIHLHQNIVDVQLYGVGSFAFVSFALKWEYIL	720
Db	661	ELSPLLSTTEWQILPCAFTLTPALSTGLIHLHQNIVDVQLYGVGSFAFVSFALKWEYIL	720
Qy	721	LLFLLADARVACILWMLLIAQAEALLENVLNNAASVAGAHGILSFLVFFCAAWYIKG	780
Db	721	LLFLLADARVACILWMLLIAQAEALLENVLNNAASVAGAHGILSFLVFFCAAWYIKG	780
Qy	781	RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCAGAVLVGLVFLTLSPYKVFILT	840
Db	781	RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCAGAVLVGLVFLTLSPYKVFILT	840
Qy	841	RLIWLQYFTRAEAHMQVWVPLNVGRGRDAIILLTCAVHPELIFDITKLLALLGLPLM	900
Db	841	RLIWLQYFTRAEAHMQVWVPLNVGRGRDAIILLTCAVHPELIFDITKLLALLGLPLM	900
Qy	901	VLOAGITRVPYFVRAQGLIRACMLVRKVAGHYQVQVPMKLGALTGTVVYNHLPRLDWA	960
Db	901	VLOAGITRVPYFVRAQGLIRACMLVRKVAGHYQVQVPMKLGALTGTVVYNHLPRLDWA	960
Qy	961	HAGRLDLAVAVEPVVFSAMETKVITWADTAAACGDIILGLPVSAARRGKEIFLGPADSLEG	1020
Db	961	HAGRLDLAVAVEPVVFSAMETKVITWADTAAACGDIILGLPVSAARRGKEIFLGPADSLEG	1020
Qy	1021	QGWRLAPITAYSQOQTRGVLCIITSLTGRDNQKQVEGEVQVVSSTATQSFATCINGVCWT	1080
Db	1021	QGWRLAPITAYSQOQTRGVLCIITSLTGRDNQKQVEGEVQVVSSTATQSFATCINGVCWT	1080
Qy	1081	VYHAGSKTLAGKPGPIQMTYTNVDLVLVQWQAPGARSMTPCSGSDLYLVTRHADVI	1140
Db	1081	VYHAGSKTLAGKPGPIQMTYTNVDLVLVQWQAPGARSMTPCSGSDLYLVTRHADVI	1140
Qy	1141	PVRRGDSRGSLLSPRPVSYLKGSGGPLLCPSHGVGVVFRAAVCTRGVAKAVDFIPVES	1200
Db	1141	PVRRGDSRGSLLSPRPVSYLKGSGGPLLCPSHGVGVVFRAAVCTRGVAKAVDFIPVES	1200
Qy	1201	METTWRSPVFTDNTSPPAVPOTFQVAHLHAPATGSGKSTKPAAYAAQGYKVLVNSVAA	1260
Db	1201	METTWRSPVFTDNTSPPAVPOTFQVAHLHAPATGSGKSTKPAAYAAQGYKVLVNSVAA	1260
Qy	1261	TLGFGAYMSKAHGIDPNIRTVRTITTTGGSITYSTYKFLADGGCGSGGAYDIIICDECHS	1320
Db	1261	TLGFGAYMSKAHGIDPNIRTVRTITTTGGSITYSTYKFLADGGCGSGGAYDIIICDECHS	1320
Qy	1321	TDSSTILGTIVLDQAETAGARLVLATATPPGCVTVPHNIBEEIGLSNNGEIPFYGKAI	1380
Db	1321	TDSSTILGTIVLDQAETAGARLVLATATPPGCVTVPHNIBEEIGLSNNGEIPFYGKAI	1380
Qy	1381	PTEAIKGGRHLLFCHSKKKDELAALKJGLGNNAVAYYRGLDVSVPPIGDVVVVAITDAL	1440
Db	1381	PTEAIKGGRHLLFCHSKKKDELAALKJGLGNNAVAYYRGLDVSVPPIGDVVVVAITDAL	1440
Qy	1441	MTGFTGDFSVIDCNTCVTQTVDFSLDPTFTTITTTVPQDAVRSQRRGTRGRSGIYR	1500
Db	1441	MTGFTGDFSVIDCNTCVTQTVDFSLDPTFTTITTTVPQDAVRSQRRGTRGRSGIYR	1500
Qy	1501	FVTPGERPSGMFSDSVLCECYDAGCAWYELTPAETSVRLRAYLNTPGLPVCQDHLEFWS	1560
Db	1501	FVTPGERPSGMFSDSVLCECYDAGCAWYELTPAETSVRLRAYLNTPGLPVCQDHLEFWS	1560
Qy	1561	VFTGLTHIDAHFISQTKQAGDNFPYLVAQYATCARAAQAPPSWQDMWCKLIRLKPTLHG	1620
Db	1561	VFTGLTHIDAHFISQTKQAGDNFPYLVAQYATCARAAQAPPSWQDMWCKLIRLKPTLHG	1620
Qy	1621	PTLLVRLGAVQNEVILTHPTIKYIMACMSADLEVTSTWVLVGGVLAALAAVCLTTGVS	1680
Db	1621	PTLLVRLGAVQNEVILTHPTIKYIMACMSADLEVTSTWVLVGGVLAALAAVCLTTGVS	1680
Qy	1681	VIVGRILISGKPAVVPDREVLYQEFDEMEECASQLPYIEQGMQLAEQFKOKALGLLAT	1740
Db	1681	VIVGRILISGKPAVVPDREVLYQEFDEMEECASQLPYIEQGMQLAEQFKOKALGLLAT	1740
Qy	1741	KQAEAAAPVSVESKWRALETFPAKHMNFIISGIVLAGLSTLPGNPATASLMAFTASITSP	1800
Db	1741	KQAEAAAPVSVESKWRALETFPAKHMNFIISGIVLAGLSTLPGNPATASLMAFTASITSP	1800
Qy	1801	LTQNTLLFNILGWNAAQALAPSAASAFVAGTAGAAGVSGIGLVKLVLDLAGYGAGVA	1860
Db	1801	LTQNTLLFNILGWNAAQALAPSAASAFVAGTAGAAGVSGIGLVKLVLDLAGYGAGVA	1860
Qy	1861	GALVAFKMSGEVSTEDLVNLLPAIILSPGALVGVVCAAILRRHRVGPGEVQVMMNRLI	1920
Db	1861	GALVAFKMSGEVSTEDLVNLLPAIILSPGALVGVVCAAILRRHRVGPGEVQVMMNRLI	1920
Qy	1921	AFASRGNHVSPTHYVVPESDAAAARVTQILSSLTITOLLKRLHQNEDCSTPCSGSWLRDV	1980
Db	1921	AFASRGNHVSPTHYVVPESDAAAARVTQILSSLTITOLLKRLHQNEDCSTPCSGSWLRDV	1980
Qy	1981	WDWTCTVLTDPKTLQSKLLPRLPGVFPFLSCORGKGVWRGDIQMOITTCGCAQIAGHVK	2040
Db	1981	WDWTCTVLTDPKTLQSKLLPRLPGVFPFLSCORGKGVWRGDIQMOITTCGCAQIAGHVK	2040
Qy	2041	NGSMRIYGPRTCSNTHGTPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDH	2100
Db	2041	NGSMRIYGPRTCSNTHGTPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDH	2100
Qy	2101	YVTGMTDNVCKPCQVPAPEFTEVDGVRLHRYAPACKPELLREDVTFQVGLNQYLVSQ	2160
Db	2101	YVTGMTDNVCKPCQVPAPEFTEVDGVRLHRYAPACKPELLREDVTFQVGLNQYLVSQ	2160
Qy	2161	PCEPEPDVTLTSMLTDPSSHITATKRLRARGSPSLASSASQLSAPSKATCTTHD	2220
Db	2161	PCEPEPDVTLTSMLTDPSSHITATKRLRARGSPSLASSASQLSAPSKATCTTHD	2220
Qy	2221	SPDADLIEANLLWROEMGNITRVESENKVVILDSFELHAEGREREISVAEILKRSK	2280
Db	2221	SPDADLIEANLLWROEMGNITRVESENKVVILDSFELHAEGREREISVAEILKRSK	2280

QY 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTSNV 2340
DB 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTSNV 2340
QY 2341 SSALAELATKTFSSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPLEGEFGDPL 2400
DB 2341 SSALAELATKTFSSGSSAVDSGTATALPDQASDDGDKGSDVESYSSMPLEGEFGDPL 2400
QY 2401 SDGSSWTVSEASESDVCCSMSTYTWGALTITCAABESKLPINPLNSLLRHHNNVYATT 2460
DB 2401 SDGSSWTVSEASESDVCCSMSTYTWGALTITCAABESKLPINPLNSLLRHHNNVYATT 2460
QY 2461 SPSASLRQKKVTFDRLOVLDDHRYDLVKEMKAKASTVAKALLSIEACKLTPPHSAKSF 2520
DB 2461 SPSASLRQKKVTFDRLOVLDDHRYDLVKEMKAKASTVAKALLSIEACKLTPPHSAKSF 2520
QY 2521 GYGAKDVRLNSRAVNHRSVWEDLLEDTEPIDITTIMAKSVFVCQPEKGGRKPARLIV 2580
DB 2521 GYGAKDVRLNSRAVNHRSVWEDLLEDTEPIDITTIMAKSVFVCQPEKGGRKPARLIV 2580
QY 2581 FPDLGVRVCEKMAVDVSTLPOAVMGSSYGFQYSPKQKRVFLVNTWKSCKCPMGFSYDT 2640
DB 2581 FPDLGVRVCEKMAVDVSTLPOAVMGSSYGFQYSPKQKRVFLVNTWKSCKCPMGFSYDT 2640
QY 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGYRRCRA 2700
DB 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGYRRCRA 2700
QY 2701 SGVLTTSCNTLITCVLKATAACRAKLDQCTMLVNGDDLWVICSAGTQEDAAALRAFTE 2760
DB 2701 SGVLTTSCNTLITCVLKATAACRAKLDQCTMLVNGDDLWVICSAGTQEDAAALRAFTE 2760
QY 2761 AMTRYSAAPGDPQPEYDLELITSCSNVSVVAHDASGKRVYVLTTRDPTPLARAWEETAR 2820
DB 2761 AMTRYSAAPGDPQPEYDLELITSCSNVSVVAHDASGKRVYVLTTRDPTPLARAWEETAR 2820
QY 2821 HTPINSWLGNIIMYAPALWARMILMTHFFSIIILAOBQLEKALDCQIYGACYSIEPLDLPQ 2880
DB 2821 HTPINSWLGNIIMYAPALWARMILMTHFFSIIILAOBQLEKALDCQIYGACYSIEPLDLPQ 2880
QY 2881 IIERLHGLSAFTLHYSYSPGEINRVASCLRKLGVPPLRTWRHRARSVRKLLSQGGRAATC 2940
DB 2881 IIERLHGLSAFTLHYSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRKLLSQGGRAATC 2940
QY 2941 GRYLFNWAVRTKLTPIPAASQDLGSMFVAGYSGGDIYHLSRARPRWFPLCULLLSV 3000
DB 2941 GRYLFNWAVRTKLTPIPAASQDLGSMFVAGYSGGDIYHLSRARPRWFPLCULLLSV 3000
QY 3001 GVGIVLLPNR 3010
DB 3001 GVGIVLLPNR 3010

RESULT 7

ADP88597 ID ADF88597 standard; protein; 3010 AA.

AC ADF88597;

DT 26-FEB-2004 (first entry)

DE Hepatitis C virus NS3 gene protein, SEQ ID No 6.

KW liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity.

OS Hepatitis C virus.

PN JP2003210181-A.

PD 29-JUL-2003.

PP 30-MAY-2002; 2002JP-00158335.

XX 16-NOV-2001; 2001JP-00352443.
XX (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.
XX WPI; 2003-819836/77.
XX N-PSDB; ADF88596.
PT Diagnosing liver cancer, involves amplifying amino terminal region of
PT hepatitis-C virus gene using predetermined primer and determining
XX hepatitis-C virus in base sequence of obtained DNA fragment.
PS Disclosure; SEQ ID NO 6; 36pp; Japanese.
XX The invention relates to the novel testing method for diagnosing liver
CC cancer. The novel method comprises amplifying the amino terminal region
CC of a hepatitis-C virus NS3 gene using a predetermined primer and
CC determining the hepatitis-C virus in a base sequence of the obtained DNA
CC fragment. The novel testing method is useful for diagnosing liver cancer
CC and also used in a gene amplification technique, a clinical laboratory
CC test reagent, a polymerase chain reaction, a base sequence analysis and
CC genetic engineering. The method enables the detection of a hepatitis-C
CC virus having high carcinogenicity with high specificity. This sequence
CC represents the protein of the hepatitis-C virus NS3 gene of the
XX invention.
XX Sequence 3010 AA;
SQ Query Match 96.0%; Score 15367; DB 7; Length 3010;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2866; Conservative 64; Mismatches 80; Indels 0; Gaps 0;
QY 1 MSTNPKPORKTKRNTNRRPQDVKPGCGQIVGGVYLLPRRGPRLGVRATRKASERSQPG 60
DB 1 MSTNPKPORKTKRNTNRRPQDVKPGCGQIVGGVYLLPRRGPRLGVRATRKTSERSQPG 60
QY 61 RROQIPKARRPEGRAWAOPGYWPPLYGNEGGLGWAQWLLSPRGSRSPSWGPTDPRRSRNLG 120
DB 61 RROQIPKARRPEGRTWAOPGYWPPLYGNEGGMWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
QY 121 KVIDLTCGFADLMGYIPLVGAPLGGAAALAHGVYRVLEDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDLTCGFADLMGYIPLVGAPLGGAAALAHGVYRVLEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIIPASAYEVNVSIGIYHVTNDCNSSIIVYEAADVIMHTPGCVPCVQCSNRCVW 240
DB 181 LLSCLTIIPASAYEVNVSIGIYHVTNDCNSSIIVYEAADVIMHTPGCVPCVQCSNRCVW 240
QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYVGLCGSIFLVLSQFTFSPRRHET 300
DB 241 ALTPTLAARNSSIPTTIRRHVDLLVGAALCSAMVYVGLCGSVFLVSQFTFSPRYET 300
QY 301 VQDCNCSIYPGHVSGHRMAWDMMMWNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGLA 360
DB 301 VQDCNCSIYPGHVSGHRMAWDMMMWNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGLA 360
QY 361 YYSVMGNWAKULIALLFAGVDGETHTTGRVAGHTTSGFTSLFSSCASOKIQLVNTNGSW 420
DB 361 YYSVMGNWAKULIALLFAGVDGHTHTVGRVASSTQSLVSVLWSQGPSQKIQLVNTNGSW 420
QY 421 HINRTALNCNDSLOTGFFAALFYAHKFNSGCGPERMASCRPIDWFAQGPGPIITYKPNSS 480
DB 421 HINRTALNCNDSLOTGFFAALFYAHFRNASGCGPERMASCRPIDWFAQGPGPIITHDPNSS 480
QY 481 DORPYCWHYAPPCGVVPASQVCGPVYCFPTSPVVVGTTRDSGVPTYSWGENETDVMLLN 540
DB 481 DORPYCWHYAPPCGVVPASQVCGPVYCFPTSPVVVGTTRDREGAPTYSWGENETDVLSS 540
QY 541 NTRPPQGNWFGCTWMNSTGFTTKTCGPPCNICGVGNRTLIPTDCFRKHPKPEATYTCGSG 600
DB 541 NTRPPQGNWFGCTWMNSTGFTTKTCGPPCNICGVGNRTLIPTDCFRKHPKPEATYTCGSG 600
QY 601 PWLTPRCLVDYPRYLRWHVPCTLNFISIFKVRMYVGVGVEHRLNAAACNWRGRCNLEDRDRS 660

Db 601 PMLTPRCWVDYVYRLWHYPCVTNFTVKVRMYVGGVEHRLNAACNWRGRCDEDRDRS 660
Qy 661 ELSPLLSSTTEWQILPCAETTLIPALSTGLIHHLHONIVDVQYLYGVCSAFVSFAIKWEYIL 720
Db 661 ELSPLLSSTTEWQILPCSTFTPLPALSTGLIHLHNRNIVDVQYLYGVCSAVVSFAIKWEYIL 720
Qy 721 LFLFLLLADARVCAACLMWMLLIAQAEALLENLVVINAASVAGAHGILSFLVFFCAAWYIKG 780
Db 721 LFLFLLLADARVCAACLMWMLLIAQAEATLENLVVINAASVAGAHGILSFLVFFCAAWYIKG 780
Qy 781 RLAPGAAYAFYGVWPILLLLLALPPRAYALDRMAASCGAVLVGLVFLTSLSPYKVFILT 840
Db 781 RLVPGAAYALYGVWPILLLLLALPPRAYAMDRMAASCGAVFVGLVLLTSLSPYKVFLA 840
Qy 841 RLILWMLQYFTRAEAHQWVPLNVVRGGRDAIILLTCAVHPELIPDITKLLAAILGPLM 900
Db 841 RLILWMLQYFTRAEAHQWVPLNVVRGGRDAIILLTCAVHPELIPDITKLLAAILGPLM 900
Qy 901 VLOAGITRVYFVRAOGLIRACMLVRKVAGHYVQVMFKLGALITGTYYVYNHLTPLRDWA 960
Db 901 VLOAGITRVYFVRAOGLIRACMLVRKVAGHYVQVMFKLAALITGTYYVYVYDHLTPLRDWA 960
Qy 961 HAGLRDLAVAVEPVVFSAMETKVTWGDATTAACGDIILGLPVSARRKEIFLGPADSLEG 1020
Db 961 HAGLRDLAVAVEPVVFSMETKLTITWGADTAACGDIISGLPVSARRKEIILLGPADSFE 1020
Qy 1021 QGWRLLAPIITAYSQOTRGVLCIITSITGRDKNOVEGEVQVVSSTATOSFLATCINGVCWT 1080
Db 1021 QGWRLLAPIITAYSQOTRGVLCIITSITGRDKNOQVGEVQVVSSTATOSFLATCINGVCWT 1080
Qy 1081 VYHGAGSKTLAGPKPITQMYTNVDLGVQWQAPPGARSMTPCSCGSDLYLVTRHADVI 1140
Db 1081 VYHGAGSKTLAGPKPITQMYTNVDQDLVWPAPPGARSMTPTCTCGSSDLYLVTRHADVV 1140
Qy 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGLPCPSGHVYGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRRGDSRGSLLSPRPISYLYKGSSGGLPCPSGHVYGVIFRAAVCTRGVAKAVDFIPVES 1200
Qy 1201 METTMRSPVFDNSTPAPVOTFOVAHLHAPTGSKSTKVPAAVAAOQYKVLVLPNSVAA 1260
Db 1201 METTMRSPVFDNSTSPAPVOTFOVAHLHAPTGSKSTKVPAAVAAOQYKVLVLPNSVAA 1260
Qy 1261 TLGFGAYMSKAHGIDPNIRTCVRITTTGGSITYSTYKFLADGCGSGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGIEPNIRTCVRITTTGGPITYSTYKFLADGCGSGAYDIIICDECHS 1320
Qy 1321 TDSITTLIGITVLDOAETAGARLVVLATATPPGSVTVPHPNIEEIGLSNNGEIPFYGKAI 1380
Db 1321 TDSITTLIGITVLDOAETAGARLVVLATATPPGSITVPHPNIEEVALSNTGEIPFYGKAI 1380
Qy 1381 PTEAIKGGRHILFCHSKKCCDELAALTGGLGNNAVAYYRGLDVSVIPIPGDVVVATDAL 1440
Db 1381 PTEAIKGGRHILFCHSKKCCDELAALTGGLGNNAVAYYRGLDVSVIPTSGDVVVVATDAL 1440
Qy 1441 MTGFTGDFSDVDCNCTQTVDFSLDPTFTIETITVPDQAVSRORGRGSGIYR 1500
Db 1441 MTGFTGDFSDVDCNCTQTVDFSLDPTFTIETITLPDQAVSRQORGRGSGIYR 1500
Qy 1501 FVTPCGRSGMFDSSVLCBYDAGCAWYELTPAETSVRILRAYLNTPLGPVQCDHLEFWES 1560
Db 1501 FVTPCGRSGMFDSSVLCBYDAGCAWYELTPAETSVRILRAYLNTPLGPVQCDHLEFWES 1560
Qy 1561 VFTGLTHIDAHFSLQTKAGDNFPYLVAQATVCARAQAPPSQWMMKCLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFSLQTKAGDNLPYLVAQATVCARAQAPPSQWMMKCLIRLKPTLHG 1620
Qy 1621 PTPLLYRLGAVONEVILTHPIITYKIMACSADLEVVSTWVLVGGVLAALAYCLITGVS 1680
Db 1621 PTPLLYRLGAVONEVILTHPIITYKIMACSADLEVVSTWVLVGGVLAALAYCLITGVS 1680
Qy 1681 VLVGRILISGKPAVDPREVLYQEPDEMEECASQLPYIEQGMQLABEQFKQKALGLLQAT 1740

Db 1681 VLVGRILISGKPAVDPREVLYQEPDEMEECASHLPYIEQGMQLABEQFKQKALGLLQAT 1740
Qy 1741 KOAEAAAPVVEKSWRALETFWAKHMWNFISGTOYLAGLSTLPCNPAIASLMFTAISITSP 1800
Db 1741 KOAEAAAPVVEKSWRALETFWAKHMWNFISGTOYLAGLSTLPCNPAIASLMFTAISITSP 1800
Qy 1801 LTTQNTLLFNIIIGGWVAAQOLAPPSAASAFVAGIAGAAVGSIGLKVLDVILLAGYGAGVA 1860
Db 1801 LTTQNTLLFNIIIGGWVAAQOLAPPSAASAFVAGIAGAAVGSIGLKVLDVILLAGYGAGVA 1860
Qy 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNNRLI 1920
Db 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNNRLI 1920
Qy 1921 APASRGNHVSPTHYVPESDAAAARVTOILSSLTITQLKRLHOWINEDCSTPCSGSMLRDV 1980
Db 1921 APASRGNHVSPTHYVPESDAAAARVTOILSSLTITQLKRLHOWINEDCSTPCSGSMLKDV 1980
Qy 1981 WDWICTVLTDFTKWLQSKLLPRLPGVPFLSCORGYKGVWRGDMGIMOTTCPCGAQIAGHVK 2040
Db 1981 WDWICTVLSDFTKWLQSKLLPRLPGVPFLSCORGYKGVWRGDMGIMOTTCPCGAQITGHVK 2040
Qy 2041 NGSMRIVGPRTCSNTHWGTFFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDFH 2100
Db 2041 NGSMRIVGPRTCSNTHWGTFFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDFH 2100
Qy 2101 YVTGMTDNNVKPCQVPAPPEFFTEVDGVRILHRVAPACKPILLREDVTFQVGLNQLVGSOL 2160
Db 2101 YVTGMTDNNVKPCQVPAPPEFFTEVDGVRILHRVAPACKPILLREDVTFQVGLNQLVGSOL 2160
Qy 2161 PCPEPEPDVTLTSMITDPSHITAEAKRLARGSPPSLASSASQASAPSLKATCTTHHD 2220
Db 2161 PCPEPEPDVTLTSMITDPSHITAEAKRLARGSPPSLASSASQASAPSLKATCTTHHD 2220
Qy 2221 SPADLIEANLLWRQEMGNIITRVESENKVILDSFELPFAEGDEREISVAAEILRKSXK 2280
Db 2221 SPADLIEANLLWRQEMGNIITRVESENKVILDSFDPPIRAVEDEREISVPAEILRKPXK 2280
Qy 2281 FPSALPIWARPDYNNPILLESWKDPDYVPPVHVGCPLPPTKAPPIPPRRKRKTVVLTESNV 2340
Db 2281 FPSALPIWARPDYNNPILLESWKDPDYVPPVHVGCPLPSTKAPPIPPRRKRKTVVLTSTV 2340
Qy 2341 SSALAEALATKTFGSSGSAVDSTATATPDLASDDGDKGSDVESYSSMPPLEGPGDPL 2400
Db 2341 SSALAEALATKTFGSSGSAVDSTATGPDPQASDDGDKGSDVESYSSMPPLEGPGDPL 2400
Qy 2401 SDGSMSTVSEASDDVCCSMSTYWTGALITPCAAEESKLPINPLSNLSLRHNNWYATT 2460
Db 2401 SDGSMSTVSGAGEDVCCSMSTYWTGALITPCAAEESKLPINPLSNLSLRHNSWYSTT 2460
Qy 2461 SRSASLRQKQVTFDRLOVLDHYRDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
Db 2461 SRSASLRQKQVTFDRLOVLDHYRDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
Qy 2521 GYGAKDVNRNLSRAVNHIRSWMEDLLETETPDIITIMAKSEVFCVQEKGRKPARLIV 2580
Db 2521 GYGAKDVNRNLSRAVNHIRSWMEDLLETETPDIITIMAKSEVFCVQEKGRKPARLIV 2580
Qy 2581 FPDIGVRVCEKMAKYDVVSTLPOAVMGSSYGFQYSPKORVEFLVNTWKSKKCPMGFSYDT 2640
Db 2581 FPDIGVRVCEKMAKYDVVSTLPOAVMGPSYGFQYSPGORVEFLVNTWKSKKCPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVEESYQCCDLAPEARQAIRSLITERLYIGGPLTNSKGQNGCYRRCRA 2700
Db 2641 RCFDSTVTESDIRVEESYQCCDLAPEARQAIRSLITERLYIGGPLTNSKGQNGCYRRCRA 2700
Qy 2701 SGVLTTSCTNTLTCYLKATAACRAAKLQDCTMLVNGDDLWICESAGTOEDAAALRAFTE 2760
Db 2701 SGVLTTSCTNTLTCYLKATAACRAAKLQDCTMLVNGDDLWICESAGTOEDAAALRAFTE 2760
Qy 2761 AMTRYSAAPPDPPQPEYDELELITSCSNVSVAHADSGKRVVYLTDRDPTTPLARAAWETAR 2820
Db 2761 AMTRYSAAPPDPPQPEYDELELITSCSNVSVAHADSGKRVVYLTDRDPTTPLARAAWETVR 2820

```
QY 2821 HTPINSWLGNIIMYAPTLWARMILMTHPFSIILAOEQLKALDCOYGACYSIEPLDLPO 2880
DB 2821 HTPVNSWLGNIIMYAPTLWARMILMTHPFSIILAOEQLKALDCOYGACYSIEPLDLPO 2880
QY 2881 IIERLHGLSAFTLHYSYSGEINRVASCLRKLGVPPLRTWRHRARSVRAKLLSQGGAATC 2940
DB 2881 IIERLHGLSAFTLHYSYSGEINRVASCLRKLGVPPLRVHRARSVRAKLLSQGGAATC 2940
QY 2941 GRYLENWAVRTKLTPIPAASOLDLSGWFVAGYSGGDIYHLSRARPRWPLCLILLSV 3000
DB 2941 GRYLENWAVRTKLTPIPAASOLDLSGWFVAGYSGGDIYHLSRARPRWPLCLILLSV 3000
QY 3001 GVGIVLLPNR 3010
DB 3001 GVGIVLLPNR 3010

RESULT 8
ID ADX40808 standard; protein; 3010 AA.
XX AC ADX40808;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #31.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX DR
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX PS Disclosure; Page 388-440; 458pp; English.
XX CC
XX CC The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;

Query Match 96.0%; Score 15367; DB 9; Length 3010;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2866; Conservative 64; Mismatches 80; Indels 0; Gaps 0;

QY 1 MSTNPKQKTKRNTNRPODKVFGGGQIVGGVYLLPRGRGLGYRATRKASERSQPRG 60
DB 1 MSTNPKQKTKRNTNRPODKVFGGGQIVGGVYLLPRGRGLGYRATRKTSERSQPRG 60
QY 61 RRQPIKARRPEGRAWAQPCYPWPLYCNEGLGWAGWLLSPGRSRPSWGPTDPRRSRNIG 120
```

```
DB 61 RRQPIKARRPEGRWAQPGYFWPLYGNEGMWAGWLLSPGRSRPSWGPTDPRRSRNIG 120
QY 121 KVIDTLTCGFDLACYIPLVGAPlGGAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFDLACYIPLVGAPlGGAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIPASAYEVNRVNSGIYHVNTDCSNSSIVYEAAADVIMHTPGCVPCQVQSGNSRCWV 240
DB 181 LLSCLTIPASAYEVNRVNSGIYHVNTDCSNSSIVYEAAADVIMHTPGCVPCVRSNRCWV 240
QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVGDLGSIIFLVSQLFTFSRRRHET 300
DB 241 ALTPTLAARNSSIPPTTIRRHVDLLVGAALCSAMVVGDLGCVFLVLSOLFSTSPRRYET 300
QY 301 VODCNCSIYPGHVSGHRNADMMNNWSPPTALVVSQLLRIPOAVVDMVAGAHGVLGALA 360
DB 301 VODCNCSIYPGHVSGHRNADMMNNWSPPTALVVSQLLRIPOAVVDMVAGAHGVLGALA 360
QY 361 YYSVMGNWAKVLIIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASOKIOLVNTNGSW 420
DB 361 YYSVMGNWAKVLIIVMLLFAGVDGHTHTVGRVASTQSILVSWLSQGPSQKIQLVNTNGSW 420
QY 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCGPERMASCRPIDWFAQGMPITYTKPNSS 480
DB 421 HINRTALNCNDSLOTGFFAALFYAHFNASGCGPERMASCRPIDWFAQGMPITHDMPSS 480
QY 481 DORPYCWHYARPCGVPASQVCGPVYCFTPSPVVVGTTRDSGVPTYSNGENETDVLNLN 540
DB 481 DORPYCWHYARPCGIVPASQVCGPVYCFTPSPVVVGTTRDFGAPTYSGENETDVLNLS 540
QY 541 NTRPPQGNWFGCTWNNSTGFTKTCGPPCNIGVGNRTLICPTDCFRKHPEATYTKCGSG 600
DB 541 NTRPPQGNWFGCTWNNSTGFTKTCGPPCNIGVGNNTLVCFDTCFRKHPEATYTKCGSG 600
QY 601 PMLTPRCLVDYPRYRLWHYPCTLNFSIFKVRMYVGVGVEHRLNAAACNWTGERCNLEDRDS 660
DB 601 PMLTPRCWVDYPRYRLWHYPCTVNFVFKVRMYVGVGVEHRLNAAACNWTGERCNLEDRDS 660
QY 661 ELSPLLLSTTEWQIILPCAFITLIPALSTGLIHLHQNIVDQYLYGVGSAPVSAIKWEYIL 720
DB 661 ELSPLLLSTTEWQIILPCFTLIPALSTGLIHLHNRNIVDQYLYGIGSAVVSFAIKWEYIL 720
QY 721 LFLLLADARVCACLMWMLLIQAEEALENLVNLNAAASVAGAHGILSLVFFCAAWYIKG 780
DB 721 LFLLLADARVCACLMWMLLIQAEEATLENLVNLNAAASVAGAHGILSLVFFCAAWYIKG 780
QY 781 RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCAGVLVGLVFLTLSPYKVFUT 840
DB 781 RLVPGAAYALYGVWPLLLLLALPPRAYAMDREMAASCAGVAVGLVLLTLSPYKVFULA 840
QY 841 RLIIWVLOVFIPTAEAHMQVWVPLNVGRGDDAIILLTCVAHPHELIFDIITKLIIILGLPM 900
DB 841 RLIIWVLOVFIPTAEAHMQVWVPLNVGRGDDAIILLTCVAHPHELIFDIITKLIIILGLPM 900
QY 901 VLOAGITRVFVRAQGLIRACMLVRKVAGHYVQVFMKLGALGTYYVYNHLTPLRDWA 960
DB 901 VLOAGITRVFVRAQGLIRACMLVRKVAGHYVQVFMKLAALGTYYVYDHLTPLRDWA 960
QY 961 HAGLRDLAVAVPPVFSAMETKVIITWADTAACGDIILGLPVYSARRGKEIFLGPADSLEG 1020
DB 961 HAGLRDLAVAVPPVFSDMETKLIITWADTAACGDIISGLPVYSARRGKEILLGPADSFG 1020
QY 1021 QWRLLAPITAYSQOTRGVLGCIITSLTRGDKNVEGEVQVSTATQSFATCINGVCWT 1080
DB 1021 QWRLLAPITAYSQOTRGVLGCIITSLTRGDKNVEGEVQVSTATQSFATCINGVCWT 1080
QY 1081 VYHAGSKTLAGKGPITQMTYNTVDLVLVGWQAPGARSMTPCSCGSSDLXVTRHADVI 1140
DB 1081 VYHAGSKTLAGKGPITQMTYNTVDQDLVGVWAPPGARSMTPTCTCGSSDLXVTRHADVV 1140
QY 1141 PVRRGDSRGSLSPRPVSYLKGSGGPLLCSGHVGVVFRAAVCTRGVAKAVDFIPVS 1200
```


Db 1141 PVRRRGDSRGLSPRPI SYLKSSGGPLLCPSGHVVGIIFRAAVCTRGVAKAVDFIPVES 1200
Qy 1201 METTMRSPVFTDSTNPAPVQTFQVAHLHAPTSGKSTKVPAAAYAAQGYKVLVLPNSVAA 1260
Db 1201 METTMRSPVFTDSTNPAPVQTFQVAHLHAPTSGKSTKVPAAAYAAQGYKVLVLPNSVAA 1260
Qy 1261 TLGFAGYMSKAHGDIPNIRGVRTITTTGGSTTSTYTKYKFLADGGCSGAYDIIICDECHS 1320
Db 1261 TLGFAGYMSKAHGEIPNIRGVRTITTTGGPTTSTYCKFLADGGCSGAYDIIICDECHS 1320
Qy 1321 TDSITILGICVTLDQAEATAGARLVVLTATPPGVTVPHPNIBIGLSNNGEIPFYGKAI 1380
Db 1321 TDSITILGIGVLDQAEATAGARLVVLTATPPGSIITPHNIBESVALSNTGEIPFYGKAI 1380
Qy 1381 PIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVAYYRGIDVSVIPIPGDVVVVATDAL 1440
Db 1381 PIEAIKGRHLIFCHSKKKCDLAALKLTGLGLNAVAYYRGIDVSVIPTSGDVVVVATDAL 1440
Qy 1441 MTGFTGDFSVIDCNTCTVTQVDFSLDPTFTIETTVPQDAVRSORRGRTGRGRSGIYR 1500
Db 1441 MTGFTGDFSVIDCNTCTVTQVDFSLDPTFTIETTLPQDAVSRARRGRTGRGRSGIYR 1500
Qy 1501 FVTGPERPSGFMFDSVLCCEYDAGCANYELTPAETSUPLRAYINTPGLPVQCDDHLEWES 1560
Db 1501 FVTGPERPSGFMFDSVLCCEYDAGCANYELTPAETSUPLRAYINTPGLPVQCDDHLEWES 1560
Qy 1561 VFTGTIHTDAHFLSOTKQAGNFPVLVAYQATVCARAQAPPSPDQMWKCLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSOTKQAGNLPVLVAYQATVCARAQAPPSPDQMWKCLIRLKPTLHG 1620
Qy 1621 PTPLLYRLGAVQNEVILTHPIKTYIMACMSADLEVVTSWVLGVLAALAAAYCLTTGVS 1680
Db 1621 PTPLLYRLGAVQNEVILTHPIKTYIMACMSADLEVVTSWVLGVLAALAAAYCLTTGVS 1680
Qy 1681 VIVGRILILSGKPAVVDREVLVYQEPDEMEBCASQLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Db 1681 VIVGRILILSGRPVAVVDREVLVYQEPDEMEBCASHLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Qy 1741 KOAEAAAAPVBSKWALETFWAKHWNFTISGQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KOAEAAAAPVBSKWALEVFVAKHWNFTISGQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Qy 1801 LTTQNTLLFNILGCVAAQAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGCVAAQAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVA 1860
Qy 1861 GALVAFKVMGSEVPSTEDLVNLLPALSPGALVGVWCACILRRHRVGPGEVAVQWNNRLI 1920
Db 1861 GALVAFKVMGSEMPSTEDLVNLLPALSPGALVGVWCACILRRHRVGPGEVAVQWNNRLI 1920
Qy 1921 APASRGNHVSPTHYVPESDAAARVTOILSSLTITQLLKLHOWINEDCSTPCSGSMLRDV 1980
Db 1921 APASRGNHVSPTHYVPESDAAARVTOILSSLTITQLLKLHOWINEDCSTPCSGSMLRDV 1980
Qy 1981 WDWICTVLDFKTLQSLKLLPRLPGVPLSCORGKVGWVRGDMQITTCPCQAQIAGHVK 2040
Db 1981 WDWICTVLDFKTLQSLKLLPRLPGVPLSCORGKVGWVRGDMQITTCPCQAQIAGHVK 2040
Qy 2041 NGSMRIVGPRGTSNTHGTFPINAYTTGPTSPAPNYGRALMRVAAEEYVEVTRVGDPH 2100
Db 2041 NGSMRIVGPRGTSNTHGTFPINAYTTGPTSPAPNYGRALMRVAAEEYVEVTRVGDPH 2100
Qy 2101 VYTGMTTNDVKPCQVPAPEFFTEVDGVLRLHYAPACKPILLREDDVTFQVGLNQYLVSQJL 2160
Db 2101 VYTGMTTNDVKPCQVPAPEFFTEVDGVLRLHYAPACKPILLREDDVTFQVGLNQYLVSQJL 2160
Qy 2161 PCEPEPDVLTSMLTDPSSHITAEAKRLARGSPESLASSASQSLASLKATCTTHD 2220
Db 2161 PCEPEPDVLTSMLTDPSSHITAEAKRLARGSPESLASSASQSLASLKATCTTHD 2220
Qy 2221 SPDADLIEANLLWRQEWGNNITRVESKNKVLDSPEPLHAEGDEREISVAAEILKRSK 2280
Db 2221 SPDADLIEANLLWRQEWGNNITRVESKNKVLDSFDPTRAVEDEDEISVPAEILKRPK 2280

Qy 2281 FPSALPIWAPDYNPPLLESWKDPDYVPVHVHCPLPPTKAPPIPPRRKRTTVLTESNV 2340
Db 2281 FPPALPIWAPDYNPPLLESWKDPDYVPVHVHCPLPPTKAPPIPPRRKRTTVLTESNV 2340
Qy 2341 SSALAEALATKTFGSSGSSAVDSTATATPDLASDDGDKGSDVESYSMPLEGEPPDPL 2400
Db 2341 SSALAEALATKTFGSSGSSAVDSTATATPDLASDDGDKGSDVESYSMPLEGEPPDPL 2400
Qy 2401 SDGSWSTVSSEASEDVVCCSMSTYTWTGALITPCAAEESKLPINPLNSLRLHHNMVYATT 2460
Db 2401 SDGSWSTVSSEASEDVVCCSMSTYTWTGALITPCAAEESKLPINPLNSLRLHHNMVYATT 2460
Qy 2461 SRASLRQKKVTFDRLOVLDHDDHYRDLKEMKAKASTVKALLSTIEACKLTPPHSAKSKF 2520
Db 2461 SRASLRQKKVTFDRLOVLDHDDHYRDLKEMKAKASTVKALLSTIEACKLTPPHSAKSKF 2520
Qy 2521 GYGAKOVNLSRAVNHRSWEDLLEDTETPTDTTITMAKSEVFCVQPEKGRKPARLIIV 2580
Db 2521 GYGAKOVNLSRAVNHRSWEDLLEDTETPTDTTITMAKSEVFCVQPEKGRKPARLIIV 2580
Qy 2581 FPDGLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQRVFELVNTWKSCKCPMGFSYDT 2640
Db 2581 FPDGLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQRVFELVNTWKSCKCPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVEBSIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNCQYRCRA 2700
Db 2641 RCFDSTVTENDIRTEESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNCQYRCRA 2700
Qy 2701 SGVLTTSCGNLTLCYLKATAACRAAKLODCTMLVNGDDLVIICBSAGTQEDAAALRAFT 2760
Db 2701 SGVLTTSCGNLTLCYLKATAACRAAKLODCTMLVNGDDLVIICBSAGTQEDAAALRAFT 2760
Qy 2761 AMTRYSAPPDPPPOPEYDLELITSCSSNVSAHDASGKRVYVYLRDPTTPLARAAMETAR 2820
Db 2761 AMTRYSAPPDPPPOPEYDLELITSCSSNVSAHDASGKRVYVYLRDPTTPLARAAMETAR 2820
Qy 2821 HTPNSWLNIGNIIMVAPTILWARMILMTHFFSILLAOEKLADCOIYGACYSIEPLDLQP 2880
Db 2821 HTPNSWLNIGNIIMVAPTILWARMILMTHFFSILLAOEKLADCOIYGACYSIEPLDLQP 2880
Qy 2881 IIEELHGLSAPFTLHSPGGEINRVASCLRLGVPLRTWRHRARSVRKALLSQGGRAATC 2940
Db 2881 IIEELHGLSAPFTLHSPGGEINRVASCLRLGVPLRTWRHRARSVRKALLSQGGRAATC 2940
Qy 2941 GRYLFPNAVTRTKLKLTPIPAASOLDLSCGFVAGYSGGDIYHSLSRARPRWFPCLLLLSV 3000
Db 2941 GRYLFPNAVTRTKLKLTPIPAASOLDLSCGFVAGYSGGDIYHSLSRARPRWFPCLLLLSV 3000
Qy 3001 GVGIVLLPNR 3010
Db 3001 GVGIVLLPNR 3010

RESULT 9

AAR68864
ID AAR68864 standard; protein; 3010 AA.

AC AAR68864;

DT 06-DEC-1995 (first entry)

XX Hepatitis C virus RNA helicase.

XX Hepatitis C virus; HCV; non-A non-B; helicase gene; RNA helicase;
KW baculovirus; recombinant production.

XX Hepatitis C virus.

XX Key Location/Qualifiers

FT Region 196..198

FT /label= N-linked glycosylation site

FT Region 209..211

Db	1141	PVRRRGDSRGLSPRPISVYKSSGGPLLCPSGHVVGIFRAAVCTRGVAKAVDFIPVES	1200
Qy	1201	METTMRSPVTDNSTPPAVPQTQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAA	1260
Db	1201	METTMRSPVTDNSTPPAVPQTQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAA	1260
Qy	1261	TLGFGAYMSKAGIDNIRGVTITTTGGSIITVSTYKGLADGCGSGGAYDIIICDECHS	1320
Db	1261	TLGFGAYMSKAGIDNIRGVTITTTGGSIITVSTYKGLADGCGSGGAYDIIICDECHS	1320
Qy	1321	TDSITTLIGTGLVDOAETAGARLVWLATATPPGCVTPPHENIEEIGLSNNGEIPFYGKAI	1380
Db	1321	TDSITTLIGTGLVDOAETAGARLVWLATATPPGCVTPPHENIEEIGLSNNGEIPFYGKAI	1380
Qy	1381	PIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVAYYRGDVSVPPIGVVVVATDAL	1440
Db	1381	PIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAVAYYRGDVSVPPIGVVVVATDAL	1440
Qy	1441	MTGFTGDFSDVIDNCNTCVTQTVDFSLDPTTITETTVQDAVRSQRRGRTGRSGIYR	1500
Db	1441	MTGFTGDFSDVIDNCNTCVTQTVDFSLDPTTITETTVQDAVRSQRRGRTGRSGIYR	1500
Qy	1501	FVTPGERPSGWFSSVLCBCVDAGCAWYELTPAETSVRLRAYLNTPGLPVCQDHLEFWE	1560
Db	1501	FVTPGERPSGWFSSVLCBCVDAGCAWYELTPAETSVRLRAYLNTPGLPVCQDHLEFWE	1560
Qy	1561	VFTGLTHIDAHFLSQTQKQAGDNFPYLVAQYATVCARAQAPPPSWDQWKCLIRLKP	1620
Db	1561	VFTGLTHIDAHFLSQTQKQAGDNFPYLVAQYATVCARAQAPPPSWDQWKCLIRLKP	1620
Qy	1621	PTPLLYRLGAVQNEVILTHPIKTYIMACMSADLEVTSTWVLVGGVLAALAAAYCLIT	1680
Db	1621	PTPLLYRLGAVQNEVILTHPIKTYIMACMSADLEVTSTWVLVGGVLAALAAAYCLIT	1680
Qy	1681	VIYGRITLISGKPAVVPREVLYQBFDEMBECASOLPYIEOGMQLAEQFKOKALGLLOT	1740
Db	1681	VIYGRITLISGKPAVVPREVLYQBFDEMBECASOLPYIEOGMQLAEQFKOKALGLLOT	1740
Qy	1741	KQAEAAAPVSVESKRALETWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP	1800
Db	1741	KQAEAAAPVSVESKRALETWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP	1800
Qy	1801	LTTQNTLLFNILGWWAAQLAPPSAASAFVAGAGIAGAAVSGIGLKVLDVILAGYGAGVA	1860
Db	1801	LTTQNTLLFNILGWWAAQLAPPSAASAFVAGAGIAGAAVSGIGLKVLDVILAGYGAGVA	1860
Qy	1861	GALVAFKVMSEVFSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGECAVQMMNRLI	1920
Db	1861	GALVAFKVMSEVFSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGECAVQMMNRLI	1920
Qy	1921	AFASRGNHVSPTHVYPESDAAARVTOILSSLTITQLLKRLHOWINEDCSTPCSGSWLRDV	1980
Db	1921	AFASRGNHVSPTHVYPESDAAARVTOILSSLTITQLLKRLHOWINEDCSTPCSGSWLRDV	1980
Qy	1981	WDWICTVLDTFETWLQSKLLPRLPGVPLFSCQRYKGVWRGDGIMQITPCCGAQTGHVK	2040
Db	1981	WDWICTVLDTFETWLQSKLLPRLPGVPLFSCQRYKGVWRGDGIMQITPCCGAQTGHVK	2040
Qy	2041	NGSMRIVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDFH	2100
Db	2041	NGSMRIVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDFH	2100
Qy	2101	YVTGTTDNVCKPCQVPAPEFTEVDGVLRLHRYAPACKPLLRREDVTFQVGLNOYLVGSQL	2160
Db	2101	YVTGTTDNVCKPCQVPAPEFTEVDGVLRLHRYAPACKPLLRREDVTFQVGLNOYLVGSQL	2160
Qy	2161	PCEPEPDVTLTSMLTDPDSHITAEATKRLARGSPPSLASSASOLSAPSLKATCTTHD	2220
Db	2161	PCEPEPDVTLTSMLTDPDSHITAEATKRLARGSPPSLASSASOLSAPSLKATCTTHD	2220
Qy	2221	SPDADLIEANLLWROEMGNIITRVESENKVWILDSFPIRAVEDEREISVPAEILRKPRK	2280

Db	2221	SPDADLIEANLLWROEMGNIITRVESENKVWILDSFPIRAVEDEREISVPAEILRKPRK	2280
Qy	2281	FPFALPIWARPDPYNPPLLESWKDPDYPVPVHGCPPLPPTKAPPIPPRKRKRTVLTESNV	2340
Db	2281	FPFALPIWARPDPYNPPLLESWKDPDYPVPVHGCPPLPPTKAPPIPPRKRKRTVLTESNV	2340
Qy	2341	SSALAEALATKTFGSSGSSAVDGTATATPOLASDDGDKGSDVESYSSMPPELEGPDPL	2400
Db	2341	SSALAEALATKTFGSSGSSAVDGTATATPOLASDDGDKGSDVESYSSMPPELEGPDPL	2400
Qy	2401	SDGSWSTVEEASEDEVVCCSMSTYTWTGALITPCAAEESKLPINPLSNLLRHNMVYATT	2460
Db	2401	SDGSWSTVEEASEDEVVCCSMSTYTWTGALITPCAAEESKLPINPLSNLLRHNMVYATT	2460
Qy	2461	SRASLRQKVTFRDLQVLDHRYDVLKEMKAKASTVKALLSIEBACKLTPPHSAKSKF	2520
Db	2461	SRASLRQKVTFRDLQVLDHRYDVLKEMKAKASTVKALLSIEBACKLTPPHSAKSKF	2520
Qy	2521	GYGAKDVRNLSRAVNHISVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLIV	2580
Db	2521	GYGAKDVRNLSRAVNHISVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLIV	2580
Qy	2581	FPDLGVVCEKMALYDVWSTLPOAVMGSSYGFQYSPKQRFVFLVNTWKSCKPMGFSYDT	2640
Db	2581	FPDLGVVCEKMALYDVWSTLPOAVMGSSYGFQYSPKQRFVFLVNTWKSCKPMGFSYDT	2640
Qy	2641	RCFDSVTSTESDIRVESIYQCCDLAPEARQAIISLTERLYIGGPLTNSKQNCYRRCRA	2700
Db	2641	RCFDSVTSTESDIRVESIYQCCDLAPEARQAIISLTERLYIGGPLTNSKQNCYRRCRA	2700
Qy	2701	SGVLTSSCGNTLCYLKATAACRAAKLODCTMLVNGDDLVIICESAGTQEDAAALRAFTE	2760
Db	2701	SGVLTSSCGNTLCYLKATAACRAAKLODCTMLVNGDDLVIICESAGTQEDAAALRAFTE	2760
Qy	2761	AMTRYAPPDGPQPEYDLELITSCSSNVSAHDASGKRVYVYLTDRPTTPLARAAMETAR	2820
Db	2761	AMTRYAPPDGPQPEYDLELITSCSSNVSAHDASGKRVYVYLTDRPTTPLARAAMETAR	2820
Qy	2821	HTPINSWGLNIMVAPTLWARMILMTHFFSILLAOEOLKALDCQIYGACYSIEPLDLPQ	2880
Db	2821	HTPINSWGLNIMVAPTLWARMILMTHFFSILLAOEOLKALDCQIYGACYSIEPLDLPQ	2880
Qy	2881	IIERLHGLSAPTSLSHSPGSEINRVASCLRKLGVPPLRTWRHARSVRAKLLSQCGRAATC	2940
Db	2881	IIERLHGLSAPTSLSHSPGSEINRVASCLRKLGVPPLRTWRHARSVRAKLLSQCGRAATC	2940
Qy	2941	GRYLFNWAATKLTPIPAASQDLDSGWFGVAGSGGDIYHSLSRARPRFWPLCLLLSV	3000
Db	2941	GRYLFNWAATKLTPIPAASQDLDSGWFGVAGSGGDIYHSLSRARPRFWPLCLLLSV	3000
Qy	3001	GVGIYLLPNR 3010	
Db	3001	GVGIYLLPNR 3010	
RESULT 10			
AAR82694			
ID	AAR82694	standard; protein; 3010 AA.	
AC	AAR82694;		
XX	AC		
DT	16-OCT-2003 (revised)		
DT	14-NOV-1996 (first entry)		
XX	Partial HCV non-structural polyprotein.		
XX	protease; hepatitis C virus; screening; inhibitor; proteolytic;		
KW	identification; cleavage.		
OS	Hepatitis C virus; Virus.		
XX	Key	Location/Qualifiers	
FT	Protein	898..1233	

FT /note= "partial proteinase; see AAR82692"
FT 992..1907
XX /note= "partial proteinase; see AAR82693"
PN JP07184648-A.
XX
PD 25-JUL-1995.
XX
PF 05-FEB-1993; 93JP-00018854.
XX
PR 07-FEB-1992; 92JP-00022657.
PR 18-SEP-1992; 92JP-00249240.
PR 04-DEC-1992; 92JP-00325303.
XX
PA (KAEN/) KAENNO K.
PA (SUMQ) SUMITOMO METAL IND LTD.
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX
DR WPI; 1995-287962/38.
DR N-PSDB; AAT03960.
XX
PT An HCV proteinase active substance - which has activity as an anti-HCV
PT agent and can be used to screen for proteinase inhibitors.
XX
PS Disclosure; Page 39-48; 52pp; Japanese.
XX
CC The present sequence is a partial Hepatitis C Virus (HCV) polyprotein
CC from the non-structural region. Partial proteinase sequences (AAR82692-
CC 93) are contained within this sequence. The proteinases can be used as
CC anti-HCV agents. They can also be used to screen cpds. for their ability
CC to inhibit their proteolytic activity. In this way proteinase inhibitors
CC can be identified. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 3010 AA;

Query Match 95.9%; Score 15355; DB 2; Length 3010;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2864; Conservative 64; Mismatches 82; Indels 0; Gaps 0;

QY 1 MSTNPKPQKRTNRRNQDVKFGGGQIVGGVYLLPRRGLGVRAATRKASERSQRRG 60
DB 1 MSTNPKPQKRTNRRNQDVKFGGGQIVGGVYLLPRRGLGVRAATRKTSERSQRRG 60
QY 61 RROPTPKARRPGRWAQPGYWPPLYGNEGWLWAGLLSPGRSPSWGPTDPRRSRNUG 120
DB 61 RROPTPKARRPGRWAQPGYWPPLYGNEGWLWAGLLSPGRSPSWGPTDPRRSRNUG 120
QY 121 KVIDLTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGYNVATGNLPGCSFSIFLLA 180
DB 121 KVIDLTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGYNVATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIPASAEVVRNVSIGYHVNTDCNSSIIVYEAAVMHTPGCVPCQEGNSRCWV 240
DB 181 LLSCLTIPASAEVVRNVSIGYHVNTDCNSSIIVYEAAVMHTPGCVPCVRESNFSRCWV 240
QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVGDLCSIFLVLSOLFSPRRHET 300
DB 241 ALTPTLAARNSSIPPTTIRRHVDLLVGAALCSAMVVGDLCSVFLVSOLFSPRRYET 300
QY 301 VQDCNCSIYPGHVSGHRMAWDMWMSPTTALVVSQLLRIPQAVVDMVAGAHGWLAGLA 360
DB 301 VQDCNCSIYPGHVSGHRMAWDMWMSPTTALVVSQLLRIPQAVVDMVAGAHGWLAGLA 360
QY 361 YYSVMGNWAKVILVALLPFAVGDETHHTGRVAGHTTSFGTSLFSSGASOKIQLVNTNGSW 420
DB 361 YYSVMGNWAKVILVALLPFAVGDETHHTGRVAGHTTSFGTSLFSSGASOKIQLVNTNGSW 420
QY 421 HINRTALNCNDSLOTGFFAALFYAHKFNSGCCPERNASCRPIDWFAQGGPITYTKPNSS 480
DB 421 HINRTALNCNDSLOTGFIALLFYAHKFNSGCCPERNASCRPIDWFAQGGPITYTKPNSS 480
QY 481 DORPYCWHYAPRCGVVPASQVCGPVYCTFTSPVVVGTGTTDRSGVPTYSWGENETDVLN 540
DB 481 DORPYCWHYAPRCGVVPASQVCGPVYCTFTSPVVVGTGTTDRSGVPTYSWGENETDVLN 540

DB 481 DORPYCWHYAPRCGVVPASQVCGPVYCTFTSPVVVGTGTTDRSGVPTYSWGENETDVLN 540
QY 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGVGNRTLICTDCFRKHPEATYTKCGSG 600
DB 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGVGNRTLICTDCFRKHPEATYTKCGSG 600
QY 601 PWTLPCLVDYPYRLWHYPCTLNFSEIFKVRMYVGGVEHRLNAACNWTGERCNLEDRDRS 660
DB 601 PWTLPCLVDYPYRLWHYPCTLNFSEIFKVRMYVGGVEHRLNAACNWTGERCNLEDRDRS 660
QY 661 ELSPLLLSTTEWQILPCAFITLPAISGLIHLHONIVDVQILYGVGSFAVSAIKWEYIL 720
DB 661 ELSPLLLSTTEWQILPCAFITLPAISGLIHLHONIVDVQILYGVGSFAVSAIKWEYIL 720
QY 721 LLEFLLLADARVCACILMMMLLIAQAEALLENVLNAASVAGAHGILSLVFPFCAAWYIKG 780
DB 721 LLEFLLLADARVCACILMMMLLIAQAEALLENVLNAASVAGAHGILSLVFPFCAAWYIKG 780
QY 781 RLAPGAAYAFYGVMPPLLLALLLPPRAYALDREMAASCAGAVLVGLVFLTSLSPYKVFIL 840
DB 781 RLAPGAAYAFYGVMPPLLLALLLPPRAYALDREMAASCAGAVLVGLVFLTSLSPYKVFIL 840
QY 841 RLIIWMLQYFITRAEHLQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGPIUM 900
DB 841 RLIIWMLQYFITRAEHLQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGPIUM 900
QY 901 VLOAGITRVPYFVRAQGLIRACMLVRKVAGGYVQVFMKLCALGTGVYVYVHLTPLRDWA 960
DB 901 VLOAGITRVPYFVRAQGLIRACMLVRKVAGGYVQVFMKLCALGTGVYVYVHLTPLRDWA 960
QY 961 HAGLRDLAVAVPVPVFSAMETKVITWADTAAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
DB 961 HAGLRDLAVAVPVPVFSAMETKVITWADTAAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
QY 1021 QWRLLAPITAYSQOTRGLVLCIITSLTRDKNQEVEGVQVSTATQSFATCINGVCWT 1080
DB 1021 QWRLLAPITAYSQOTRGLVLCIITSLTRDKNQEVEGVQVSTATQSFATCINGVCWT 1080
QY 1081 VYHGAGSKTLGAPKPIITQMYTNVDLDLVGWAQPPGARSMTPCSCGSSDLXLVTRHADVI 1140
DB 1081 VYHGAGSKTLGAPKPIITQMYTNVDLDLVGWAQPPGARSMTPCSCGSSDLXLVTRHADVI 1140
QY 1141 PVRRGDSRGSLLSPRPVSYLKGSGGGLPCPSGHVGVFRAAVCTRGVAKAVDIPVES 1200
DB 1141 PVRRGDSRGSLLSPRPVSYLKGSGGGLPCPSGHVGVFRAAVCTRGVAKAVDIPVES 1200
QY 1201 METTNRSPVFTDNSTPPAVPOTFQVAHLHAPTGSKSTKVPAAYAAQGHVVLNPSVAA 1260
DB 1201 METTNRSPVFTDNSTPPAVPOTFQVAHLHAPTGSKSTKVPAAYAAQGHVVLNPSVAA 1260
QY 1261 TLGFGAYMSKAHGDIPNRTIGVTTTTCGSIITYSTYKFLADGGCGGGAYDIIICDECHS 1320
DB 1261 TLGFGAYMSKAHGDIPNRTIGVTTTTCGSIITYSTYKFLADGGCGGGAYDIIICDECHS 1320
QY 1321 TDSITLIGITVLDQAETAGARLVVLTATPPGSVTVVPHNPTEIEGLSNNGBIFFYGKAI 1380
DB 1321 TDSITLIGITVLDQAETAGARLVVLTATPPGSVTVVPHNPTEIEGLSNNGBIFFYGKAI 1380
QY 1381 PTEALIKGRHILFCHSKKKDELAALKLGLGLNAVAYRGLDVSVIPIPTGDVVVATDAL 1440
DB 1381 PTEALIKGRHILFCHSKKKDELAALKLGLGLNAVAYRGLDVSVIPIPTGDVVVATDAL 1440
QY 1441 MTGFTGDFSDVIDCNTCTVTQVDFSLDPTFTTETTTVPQDAVSRQRRGRTGRSGIYR 1500
DB 1441 MTGFTGDFSDVIDCNTCTVTQVDFSLDPTFTTETTTVPQDAVSRQRRGRTGRSGIYR 1500
QY 1501 FVTPGERPSGMPDSSVLCEDYDAGCAWYELTTPAETSRLRAYLNTPLPGLPVCQDHLEFWES 1560
DB 1501 FVTPGERPSGMPDSSVLCEDYDAGCAWYELTTPAETSRLRAYLNTPLPGLPVCQDHLEFWES 1560
QY 1561 VFTGLTHIDAHPLSQTQAGDNFPYLVAYQATVCARAQAAPPSPWDMWKILRLKPTLHG 1620
DB 1561 VFTGLTHIDAHPLSQTQAGDNFPYLVAYQATVCARAQAAPPSPWDMWKILRLKPTLHG 1620

QY	1621	PTLLYRLGAVQNEVILTHPIITKYIIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTV	1680
DB	1621	PTLLYRLGAVQNEVILTHPIITKYIIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTV	1680
QY	1681	VIIGRIILSKPAVDPREVLYQEPDEMEECASOLPYIEGQMOLAEQFKQKALGLLQAT	1740
DB	1681	VIIGRIILSKPAVDPREVLYQEPDEMEECASOLPYIEGQMOLAEQFKQKALGLLQAT	1740
QY	1741	KQAEAAAPVVEKSWRALETFAWKHMNFISGIQYLAGLSTLPGNPAIASMAFTASITSP	1800
DB	1741	KQAEAAAPVVEKSWRALETFAWKHMNFISGIQYLAGLSTLPGNPAIASMAFTASITSP	1800
QY	1801	LTTQNTLLFNILGWAQAAPSAASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVA	1860
DB	1801	LTTQNTLLFNILGWAQAAPSAASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVA	1860
QY	1861	GALVAFKVMGSEVPSTEDVNLPAIILSPGALVGVVCAAILRRHVGPGEVQVMNRLI	1920
DB	1861	GALVAFKVMGSEVPSTEDVNLPAIILSPGALVGVVCAAILRRHVGPGEVQVMNRLI	1920
QY	1921	AFASRGNHVSPTHVPESDAAARVTQILSLITITQLLKRHLHOWNEDCSTPCSGSWLRDV	1980
DB	1921	AFASRGNHVSPTHVPESDAAARVTQILSLITITQLLKRHLHOWNEDCSTPCSGSWLRDV	1980
QY	1981	WDMICTVLTPDKTWLQSKLLPRLPGVPLFSCQGYKGVWRGDGIMQTTCPGQAIGHVK	2040
DB	1981	WDMICTVLTPDKTWLQSKLLPRLPGVPLFSCQGYKGVWRGDGIMQTTCPGQAIGHVK	2040
QY	2041	NGSMRIIVGPRCTSNWTHGTPPINAYTTGCPCTSPAPNYSRALMRVAABEYVTRVGDFF	2100
DB	2041	NGSMRIIVGPRCTSNWTHGTPPINAYTTGCPCTSPAPNYSRALMRVAABEYVTRVGDFF	2100
QY	2101	YVTGWTNDVKPCQVPAPEFTEVDGVRHRYAPACKPLLRREDDVTQVGLNQVLVGSOL	2160
DB	2101	YVTGWTNDVKPCQVPAPEFTEVDGVRHRYAPACKPLLRREDDVTQVGLNQVLVGSOL	2160
QY	2161	PCEPEPDVTLTSMLTTPSHITATAKRLARGSPPLSASSASQLSAPSLKATCTTHHD	2220
DB	2161	PCEPEPDVTLTSMLTTPSHITATAKRLARGSPPLSASSASQLSAPSLKATCTTHHD	2220
QY	2221	SPDADLIEANLLMQEWMGNITRVESENKVVILDSPEPLHAEGEREISVAABILRSRK	2280
DB	2221	SPDADLIEANLLMQEWMGNITRVESENKVVILDSFDIRAVEDEREITSVPAEILRPRK	2280
QY	2281	FPSALPIWARPDYNPPLLESWKDPDYPPVHGCPLPPTKAPPIPPRRKRTVLTESNV	2340
DB	2281	FPPALPIWARPDYNPPLLESWKDPDYPPVHGCPLPPTKAPPIPPRRKRTVLTSTV	2340
QY	2341	SSALAEATKTFGSGSSAVDSGTATAPDLASDDGKSDVESYSSMPPLEGEFGDPLD	2400
DB	2341	SSALAEATKTFGSGSSAVDSGTATGPDQASDDGKSDVESYSSMPPLEGEFGDPLD	2400
QY	2401	SDGWSVTSEASEDVCWCCSYMTWTGALITPCAABESKLPINPLNSLLRHHNNVYATT	2460
DB	2401	SDGWSVTSEAGEDVVCWCCSYMTWTGALITPCAABESKLPINPLNSLLRHHNSVYSTT	2460
QY	2461	SRSASLRQKKVTFDLRLQVLDHYRDLKEMKAKASTVKAKLLSIEACKLTPPHSAKSKF	2520
DB	2461	SRSASLRQKKVTFDLRLQVLDHYRDLKEMKAKASTVKAKLLSIEACKLTPPHSAKSKF	2520
QY	2521	GVGAKDVRNLSRAVNHRSWEDLELDETETPIDTTIMAKSEVFCVQPEKGGKRPARIIV	2580
DB	2521	GVGAKDVRNLSRAVNHRSWEDLELDETETPIDTTIMAKSEVFCVQPEKGGKRPARIIV	2580
QY	2581	FPDLGVRVCEKMAVDVWSTLTPQAVMGSSYGFQYSPKQRFVEFLVNTWKSCKCPMGFSYDT	2640
DB	2581	FPDLGVRVCEKMAVDVWSTLTPQAVMGPSYGFQYSPGQRFVEFLVNTWKSCKCPMGFSYDT	2640
QY	2641	RCFDSVTSTESDIRVESIYQCCDLAPAPARQAIRSLTERLIYIGGPLTNSKQNGCYRRCRA	2700
DB	2641	RCFDSVTSTENDIRTESIYQCCDLAPAPARQAIRSLTERLIYVGGPLTNSKQNGCYRRCRA	2700
QY	2701	SGVLTTSCGNTLTCYLKATAACRAAKLODCTMLVNGDDLWVVCESAGTOEDAAALRAFTE	2760
DB	2701	SGVLTTSCGNTLTCYLKATAACRAAKLODCTMLVNGDDLWVVCESAGTOEDAAALRAFTE	2760
QY	2761	AMTRYSAAPGDPPEYDLELITSCSSNVSVVAHDASGKRVYVLTDRPTTPLARAAWETAR	2820
DB	2761	AMTRYSAAPGDPPEYDLELITSCSSNVSVVAHDASGKRVYVLTDRPTTPLARAAWETVR	2820
QY	2821	HTPINSWLGNIIIMYAPTLWARMILMTHFFSILLAQOEKALDCQIYGACYSIEPLDLPO	2880
DB	2821	HTPVNSWLGNIIIMYAPTLWARMILMTHFFSILLAQOEKALDCQIYGACYSIEPLDLPO	2880
QY	2881	ITERLHGLSAPTLHSHYSPEINRVASCLRKLGVPPLRTWRHRARSVRAKLLSQGGAATC	2940
DB	2881	ITERLHGLSAPTLHSHYSPEINRVASCLRKLGVPPLRVHRARSVRAKLLSQGGAATC	2940
QY	2941	GRYLFNNAVTKLKTPIPAASOLDLSGMFVACVSGDGIYHSLSRARPRWFPICLLLLSV	3000
DB	2941	GYLFNNAVTKLKTPIPAASOLDLSGMFVACVSGDGIYHSLSRARPRWFMICLLLLSV	3000
QY	3001	GVGIYLLPNR 3010	
DB	3001	GVGIYLLPNR 3010	
RESULT 11			
AAR68622			
ID	AAR68622 standard; protein; 3010 AA.		
XX	AAR68622;		
AC			
XX	16-OCT-2003 (revised)		
DT	16-OCT-1995 (first entry)		
XX	HCV protein cleavable with new serine proteinase.		
DE	XX		
XX	protease; serine; cleavage; hepatitis C virus; HCV.		
KW	Hepatitis C virus; Virus.		
OS			
XX	Key Location/Qualifiers		
XX	Cleavage-site 2419..2420		
FT	/note="Serine protease cleavage site"		
XX	JP06315377-A.		
XX	15-NOV-1994.		
PD	XX		
XX	06-MAY-1993; 93JP-00105666.		
PF	XX		
XX	06-MAY-1993; 93JP-00105666.		
PR	(KAEN/) KAENNO K.		
XX	(SUMQ) SUMITOMO METAL IND LTD.		
PA	(SOYA-) SOYAKO GIJUTSU KENKYUSHO KK.		
XX	WPI; 1995-032330/05.		
DR	N-PSDB; AAQ80498.		
XX	New HCV-originated proteinase active substance - used for site-specific		
PT	cleavage by an intermolecular reaction and the purification thereof.		
XX	Disclosure; Page 10-19; 23pp; Japanese.		
PS	This protein from HCV (hepatitis C virus) (encoded by AAQ80498) is		
XX	cleaved between amino acids 2419 and 2420, by a new serine protease,		
CC	contg. the sequence of AAR6821. The proteinase is purified as a fused		
CC	product with the dihydrofolate reductase protein by using a methotrexate		
CC	column. It can be used for the development of an inhibitor for HCV		
CC	proteinase. (Updated on 16-OCT-2003 to standardise OS field)		
XX	Sequence 3010 AA;		
SQ			

Query Match																																																													
Best Local Similarity 95.8%; Score 15332; DB 2; Length 3010;																																																													
Matches 2860; Conservative 65; Mismatches 85; Indels 0; Gaps 0;																																																													
QY	1	MSTNPKPQKTKRNTNR	PDQVKFPGGQI	VGGVYLL	PRGPRLGVRAT	RKASERSQ	PRG	60																																																					
DB	1	MSTNPKPQKTKRNTNR	PDQVKFPGGQI	VGGVYLL	PRGPRLGVRAT	RKASERSQ	PRG	60																																																					
QY	61	RRQPTPKARRPEGR	AWAQPGYPWPL	YNGEGLG	WAGWLLSPRGS	RPSMGPTD	PRRRSRNLG	120																																																					
DB	61	RRQPTPKARRPEGR	TWAGQPGYPWPL	YNGEGWAGWLLSPRGS	RPSMGPTD	PRRRSRNLG	120																																																						
QY	121	KVIDTLTCG	FADLMGYIPL	VGAP	LGGAARALAH	GVRLVEDGV	NVATGNLPCG	SFIFLLA	180																																																				
DB	121	KVIDTLTCG	FADLMGYIPL	VGAP	LGGAARALAH	GVRLVEDGV	NVATGNLPCG	SFIFLLA	180																																																				
QY	181	LLSCLTTPAS	AYEVRNYS	GIYHV	TNDCSNS	SIYVEA	ADVIMHTPCG	VCPCVQEG	NSRCW	240																																																			
DB	181	LLSCLTTPAS	AYEVRNYS	GIYHV	TNDCSNS	SIYVEA	ADVIMHTPCG	VCPCVQEG	NSRCW	240																																																			
QY	241	ALTPTLAAR	NSVPTTIR	RHV	DLV	LGTAAP	CSAMYV	GDLCG	SIFLVSQ	LFTSP	RRHET	300																																																	
DB	241	ALTPTLAAR	NSVPTTIR	RHV	DLV	LGTAAL	CSAMYV	GDLCG	SFVLVSQ	LFTSP	PRYET	300																																																	
QY	301	VQDNCNSI	YPGHV	SHRM	AWMMN	WSP	TTALV	VYSQL	LRIPQ	AVD	VMVAG	HGWVLAGLA	360																																																
DB	301	VQDNCNSI	YPGHV	SHRM	AWMMN	WSP	TTALV	VYSQL	LRIPQ	AVD	VMVAG	HGWVLAGLA	360																																																
QY	361	YYSVGNW	AKULI	VALL	PAG	VDG	ETH	TTR	GVAG	HTT	SGFTS	LFSSG	ASOKIQLVNT	420																																															
DB	361	YYSVGNW	AKULI	VALL	PAG	VDG	HTH	VTG	GRV	AS	TQSLV	SWLSQ	GP	SQKIQLVNT	420																																														
QY	421	HINRTAL	CNDS	LSOT	GF	FAAL	F	YAH	KNSS	CS	CPERM	ASCR	PID	WFAO	G	GPIT	TKPNSS	480																																											
DB	421	HINRTAL	CNDS	LSOT	GF	FAAL	F	YAH	R	FN	AS	CPERM	ASCR	PID	EFAO	G	GPIT	THDMP	580																																										
QY	481	DQRPY	CW	HYA	PR	CGV	PAS	QV	CP	VPY	CF	TSP	VV	VV	GT	DRS	GYPT	YSW	GEN	ET	DM	MLN	540																																						
DB	481	DQRPY	CW	HYA	PR	CGI	V	PAS	QV	CP	VPY	CF	TSP	VV	VV	GT	DR	F	G	A	P	T	Y	SW	GEN	ET	DM	VLLS	540																																
QY	541	NTRPPQ	GNW	FG	CTW	MN	STG	FT	K	CG	PP	CNT	GG	V	GN	RT	L	CPT	DC	FR	K	H	BE	AT	Y	K	CG	SG	600																																
DB	541	NTRPPQ	GNW	FG	CTW	MN	STG	FT	K	CG	PP	CNT	GG	V	GN	NT	L	V	CPT	DC	FR	K	H	BE	AT	Y	K	CG	SG	600																															
QY	601	PWLTPR	CL	VD	YPR	LW	HY	PC	TN	LS	I	F	K	VR	Y	M	V	G	S	VE	H	R	N	A	C	N	W	T	R	GC	N	L	ED	R	D	R	S	660																							
DB	601	PWLTPR	CM	VD	YPR	LW	HY	PC	TN	LS	I	F	K	VR	Y	M	V	G	S	VE	H	R	N	A	C	N	W	T	R	GC	N	L	ED	R	D	R	S	660																							
QY	661	ELSP	LL	ST	TE	W	Q	L	P	C	A	F	T	T	L	P	A	L	S	T	G	L	I	H	L	H	O	N	I	D	V	O	Y	L	G	V	S	F	A	I	K	W	E	Y	I	L	720														
DB	661	ELSP	LL	ST	TE	W	Q	L	P	C	S	F	T	L	P	A	L	S	T	G	L	I	H	L	H	R	N	I	D	V	O	Y	L	G	I	G	S	A	V	S	F	A	I	K	W	E	Y	I	L	720											
QY	721	LF	FL	L	A	D	A	R	V	C	A	C	L	W	M	L	L	I	A	O	A	E	A	L	E	N	I	V	L	N	A	S	V	A	G	A	H	I	L	S	F	L	P	F	F	C	A	A	W	I	K	780									
DB	721	LF	FL	L	A	D	A	R	V	C	A	C	L	W	M	L	L	I	A	O	A	E	A	T	E	N	I	V	L	N	A	S	V	A	G	A	H	I	L	S	F	L	P	F	F	C	A	A	W	I	K	780									
QY	781	RLA	PC	AA	Y	F	C	W	M	P	L	L	L	L	L	A	L	P	R	A	V	A	L	D	R	E	W	A	S	C	G	G	A	V	L	C	L	V	P	L	T	S	P	Y	K	V	F	L	T	840											
DB	781	RLV	P	G	A	A	Y	F	C	W	M	P	L	L	L	L	A	L	P	R	A	V	A	L	D	R	E	W	A	S	C	G	G	A	V	F	V	G	L	V	L	L	T	S	P	Y	K	V	F	L	T	840									
QY	841	RLI	W	L	O	Y	F	I	T	R	A	E	A	H	M	O	V	P	P	L	N	V	R	G	R	D	A	I	L	L	T	C	A	V	H	P	E	L	I	F	D	I	T	K	L	L	A	I	L	G	P	M	900								
DB	841	RLI	W	L	O	Y	F	I	T	R	A	E	A	H	L	O	V	P	P	L	N	V	R	G	R	D	A	I	L	L	T	C	A	V	H	P	E	L	I	F	D	I	T	K	L	L	L	P	I	L	G	P	M	900							
QY	901	V	L	O	A	G	I	T	R	P	E	F	Y	F	V	R	A	O	G	L	I	R	A	C	M	L	V	R	K	V	A	G	G	H	Y	V	Q	M	F	M	K	G	A	L	T	G	T	Y	V	N	H	L	T	P	L	R	D	W	A	960	
DB	901	V	L	O	A	G	I	T	R	P	E	F	Y	F	V	R	A	O	G	L	I	R	A	C	M	L	V	R	K	V	A	G	G	H	Y	V	Q	M	F	M	K	L	A	L	T	G	T	Y	V	V	D	H	L	T	P	L	R	D	W	A	960
QY	961	H	A	G	L	R	D	L	A	V	A	V	E	P	F	V	S	A	M	E	T	K	I	T	W	A	D	T	A	A	C	G	D	I	L	I	G	L	P	V	S	A	R	R	G	K	E	I	F	L	G	A	D	S	L	E	G	1020			
DB	961	H	A	G	L	R	D	L	A	V	A	V	E	P	F	V	S	D	M	E	T	K	L	I	T	W	A	D	T	A	A	C	G	D	I	S	G	L	P	V	S	A	R	R	G	K	E	I	L	L	G	A	D	S	F	G	E	1020			
QY	1021	O	G	N	R	L	A	P	I	T	A	S	Q	T	R	G	L	V	G	C	I	I	T	S	L	T	R	K	N	D	K	N	O	V	E	G	V	V	V	S	T	A	T	O	S	F	L	A	T	C	I	N	G	V	C	W	T	1080			

Db 2101 YVTGMTNDVNCQCPVAPBEFFTEVDGRLHRYAPVCKLLREEVVFQVGLNQLVGSOL 2160
QY 2161 PCEPEPDVTLTSMITDPSGHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTHHD 2220
Db 2161 PCEPEPDVAVLTSMITDPSGHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTHHD 2220
QY 2221 SPADLIEANLLRWQMGNNITVESENKVILDSPEPLHAEGDEREISVAABELKSKR 2280
Db 2221 SPADLIEANLLRWQMGNNITVESENKVILDSFDPIRAVEDEREISVPABELKRPK 2280
QY 2281 FPSALPTWARPDNPPLLESWKDPDVPVPHGCPPLPPTKAPPIPPRRKRTVVLTESNV 2340
Db 2281 FPPALPIWARPDNPPLLESWKDPDVPVPHGCPPLPPTKAPPIPPRRKRTVVLTESV 2340
QY 2341 SSALAEATKTFGSSGSSAVDSGTATALPDASDDGDKGSDVESYSMPFLEGEPPDL 2400
Db 2341 SSALAEATKTFGSSGSSAVDSGTATGPPDQASDDGDKGSDVESYSMPFLEGEPPDL 2400
QY 2401 SDGSWSTVSEASEDDVCCSMSTWTGALITPCAABESKLPINPLNSLLRHNNVYATT 2460
Db 2401 SDGSWSTVSEASEDDVCCSMSTWTGALITPCAABESKLPINPLNSFLRHHSWYSTT 2460
QY 2461 SRGASLRQKVTDRLOVLDHVRDVLKEMKAKASTVKALLSIEEACKLTPPHSAKSKF 2520
Db 2461 SRGASLRQKVTDRLOVLDHVRDVLKEMKAKASTVKALLSIEEACKLTPPHSAKSKF 2520
QY 2521 GYKADVRNLSRAVNHRSVWEDLLEDTPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYKADVRNLSRAVNHRSVWEDLLEDTPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
QY 2581 FPDGLVRVCBMAVLDVWSTLPQAVMGSSYGFQSPQORVEFLVNTWKKSKCPMGFSYDT 2640
Db 2581 FPDGLVRVCBMAVLDVWSTLPQAVMGSSYGFQSPQORVEFLVNTWKKSKCPMGFSYDT 2640
QY 2641 RCFDSTVTENDIRVEESIYOCDDLAPEARQAIRSLTERLYIGGLTNSKQNGCYRRCRA 2700
Db 2641 RCFDSTVTENDIRVEESIYOCDDLAPEARQAIRSLTERLYIGGLTNSKQNGCYRRCRA 2700
QY 2701 SGVLTTSCGNLTLCYLKATACRAKLDQCTMLVNGDDLVCBSAGTQEDAAALAFTE 2760
Db 2701 SGVLTTSCGNLTLCYLKATACRAKLDQCTMLVNGDDLVCBSAGTQEDAAALAFTE 2760
QY 2761 AMTRYSAPPDPPQPEYDLELITSCSNVSVAHDSAGKRVYLLTRDPTTPLARAWEATAR 2820
Db 2761 AMTRYSAPPDPPQPEYDLELITSCSNVSVAHDSAGKRVYLLTRDPTTPLARAWEATAR 2820
QY 2821 HTPINSWLGNIIMVAPTILWARMILMTHFFSILLAOEKLDCQIYGACYSIEPLDLPQ 2880
Db 2821 HTPVNSWLGNIIMVAPTILWARMILMTHFFSILLAOEKLDCQIYGACYSIEPLDLPQ 2880
QY 2881 IIERLHGLSAFTLSHSPGEINRVASCLRLKGVPLTRHRARSVRKLLSQGGRATC 2940
Db 2881 IIERLHGLSAFTLSHSPGEINRVASCLRLKGVPLTRHRARSVRKLLSQGGRATC 2940
QY 2941 GRYLFNNAVTKLKLITPAPASQDLSCGFWVAGYSGGDIYHSLSRARPRWFLCLLLSV 3000
Db 2941 GRYLFNNAVTKLKLITPAPASQDLSCGFWVAGYSGGDIYHSLSRARPRWFLCLLLSV 3000
QY 3001 GVGIVYLLPNR 3010
Db 3001 GVGIVYLLPNR 3010

RESULT 12

ABG32451

ID ABG32451 standard; protein; 3010 AA.

XX XX

AC ABG32451;

XX 15-NOV-2002 (first entry)

DT Hepatitis C virus Con 1 isolate polyprotein.

DE XX

KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NSS; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NSSA; HCV replication; polyprotein.
XX Hepatitis C virus.
XX WO200259321-A2.
XX 01-AUG-2002.
XX 16-JAN-2002; 2002WO-EP000526.
XX 23-JAN-2001; 2001US-0263479P.
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI: 2002-599793/64.
XX N-PSDB; ABK91411.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX Claim 1; Page 34-36; 69pp; English.
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NSSA, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX ; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
XX polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
XX and NS5B proteins) used as a basis for the adaptive mutations of the
XX invention
SQ Sequence 3010 AA;
Query Match 95.7%; Score 15323; DB 5; Length 3010;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 2852; Conservative 76; Mismatches 82; Indels 0; Gaps 0;
QY 1 MSTNPKPQRTKNTNRRPODVKPPGGQIVGGVYLLPRRGPRLGVRATRKASRSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPODVKPPGGQIVGGVYLLPRRGPRLGVRATRKASRSQPRG 60
QY 61 RRQPIPKARPEGRAWAQPGYPWPPLYGNEGIGWAGLLSPRGSRPSWGTPDPRRRSRNLG 120
Db 61 RRQPIPKARPEGRAWAQPGYPWPPLYGNEGIGWAGLLSPRGSRPSWGTPDPRRRSRNLG 120
QY 121 KVDTLTGCFADLMGYIPLVGCAPLGGARALAHGVRLVEDGVNYATGNLPCCSIFILLA 180
Db 121 KVDTLTGCFADLMGYIPLVGCAPLGGARALAHGVRLVEDGVNYATGNLPCCSIFILLA 180
QY 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVVEAADVIMHTPCVPCVOEGNSRCWV 240
Db 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVVEAADVIMHTPCVPCVRENNSRCWV 240

Qy	241	ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVGDLGCSIFLVSQLFTFSRRHET	300
Db	241	ALTPTLAARNASVPTTIRRHVDLLVGAALCSAMVVGDLGCSFVLVAQLFTFSRRHET	300
Qy	301	VODCNCISYPGHVSHRMAWDMMNWSPTTALVVSQRLRIPQAVVDMVAGAHGWVLAGLA	360
Db	301	VODCNCISYPGHVSHRMAWDMMNWSPTAALVVSQRLRIPQAVVDMVAGAHGWVLAGLA	360
Qy	361	YYSMYGNNAKVLIVALLFAGVDGEHTTGRVAGHTTSGFTSLFSSGASOKIOLVNTNGSW	420
Db	361	YYSMYGNNAKVLIVMLLFAGVDGGYVTGTWAKNTLGITSLFSGSGQIKIOLVNTNGSW	420
Qy	421	HINRTALNCNDSLOQCFFAALFYAHKFNSGSGPERMASCRPDWFAQGWGPITYTKPNSS	480
Db	421	HINRTALNCNDSLNTGFLAALFYVHKFNSGSGPERMASCSPIDAFAQGWGPITYNESHS	480
Qy	481	DORPYCWHYAPRPGCVVPASQVGPVYCTPSPVVVGTTRDSGVPTYSGENETDVMLLN	540
Db	481	DORPYCWHYAPRPGCIVPAQAQVGPVYCTPSPVVVGTTRDPGVPTYSGENETDVLN	540
Qy	541	NTRPPQGNWFGCTWMNSTGFTTKCGPPCNIIGVGNRTLI CPTDCFRKHPEATYTKCGSG	600
Db	541	NTRPPQGNWFGCTWMNSTGFTTKCGPPCNIIGGINKTLTCTDCFRKHPEATYTKCGSG	600
Qy	601	PMLTPRCLVDYPRYLWHYPCTILNFSIFKVRMYGVGVEHRLNAA CNWTRGERCNLEDRDRS	660
Db	601	PMLTPRCLVHYPRYLWHYPCTVNFIFKVRMYGVGVEHLEAACNWTRGERCNLEDRDRS	660
Qy	661	ELSPILLSTTEWOIIPCAFTTLPALSTGLIHHQNIIVDQYLYGVGSAFVSPFAIKWEYIL	720
Db	661	ELSPILLSTTEWOVLPCSFTTLPALSTGLIHLHQNVVDQYLYGIGSAFVSPFAIKWEYVL	720
Qy	721	LLFLILLADARVACACLLMMLLIAQAEBALBNVYVNAASVAGAHGILSLFVFCFAAWYIKG	780
Db	721	LLFLILLADARVACACLLMMLLIAQAEBALBNVYVNAASVAGAHGILSLFVFCFAAWYIKG	780
Qy	781	RLAPGAAYAFYGVWPLLLLLLPRAVALDREMAASCAGVLGVLFITLSPYKVFIT	840
Db	781	RLVPGAAYALYGVWPLLLLLLPRAVALDREMAASCAGVAVGVGLIILTLSPHYKFLA	840
Qy	841	RLIWLQYFITRAEAHQMVVPLNVGRGRDAIILLTCAVHPBELFDITKLLAILGPILM	900
Db	841	RLIWLQYFITRAEAHLQVWIPPLNVGRGRDAVILLTCAIHPBELFITKILLAILGPILM	900
Qy	901	VLQAGITRVPYFVRAQGLIRACMLVRKVAGHYVQVFMKLGALTGTYVYNHLTPLRDA	960
Db	901	VLQAGITKVPYFVRAHGLIRACMLVRKVAGHYVQVMAKLAALTGTYVYVDHLTPLRDA	960
Qy	961	HAGLRDLAVAVBPVVSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG	1020
Db	961	HAGLRDLAVAVBPVVSFDMETKVIITWGADTAACGDIILGLPVSARRGRIHILGPADSLEG	1020
Qy	1021	QGWRLAPITAYSQOTRGLGCIITSLTGRDNQVEGEVQVYSTATQSPFLATCINGVCWT	1080
Db	1021	QGWRLAPITAYSQOTRGLGCIITSLTGRDNQVEGEVQVYSTATQSPFLATCVNGVCWT	1080
Qy	1081	VYHAGSKTLAGPKGPIITOMYTNVDLVLVQWQAPPGARGMNTPCSGSSDLYLVTRHADVI	1140
Db	1081	VYHAGSKTLAGPKGPIITOMYTNVDQLVQWQAPPGARSLTPTCCTGSSDLYLVTRHADVI	1140
Qy	1141	PVRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES	1200
Db	1141	PVRRGDSRGSLLSPRPVSYLKGSSGGPLCPSGHAAGVIFRAAVCTRGVAKAVDFVPVES	1200
Qy	1201	METMRSVPFTDNSTPPAPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLNPSVAA	1260
Db	1201	METMRSVPFTDNSSPPAPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLNPSVAA	1260
Qy	1261	TLGFGAYMSKAHGDINIRGTITGGSITYSTYKFLADGGCGSGGAYDIIICDECHS	1320
Db	1261	TLGFGAYMSKAHGDINIRGTITGAPITYSTYKFLADGGCGSGGAYDIIICDECHS	1320
Qy	1321	TDSSTILGIGTGLDQAGTARLVLTATATPPGSVTVPHPNIEEIGLSNNGNIEIPFYGRAI	1380
Db	1321	TDSSTILGIGTGLDQAGTARLVLTATATPPGSVTVPHPNIEEVALSSTGIEIPFYGRAI	1380
Qy	1381	PIEAIKGRHILFCHSKKKCDLAALKLTGLGINAVAYRGLDVSIPIPGDVVVATDAL	1440
Db	1381	PIETIKGRHILFCHSKKKCDLAALKLSGLGINAVAYRGLDVSIPTSGDVIIVATDAL	1440
Qy	1441	MTGFTGDFSDSVIDCNTCVTQVDFSLDPTFTIETTTPQDAVRSQRGRTRGRSGIYR	1500
Db	1441	MTGFTGDFSDSVIDCNTCVTQVDFSLDPTFTIETTTPQDAVRSQRGRTRGRGMIYR	1500
Qy	1501	FVTGPERPSGFMFDSVSLCECYDAGCAWYELTTPAETSRLRAYLNTPLPGLPVCODHLEFWES	1560
Db	1501	FVTGPERPSGFMFDSVSLCECYDAGCAWYELTTPAETSRLRAYLNTPLPGLPVCODHLEFWES	1560
Qy	1561	VFTGLTHIDAHPLSOTKOAGDNFPYLVAYQATVCARAQAPPPSWDMWKCLIRLKPRTLHG	1620
Db	1561	VFTGLTHIDAHFLSOTKOAGDNFPYLVAYQATVCARAQAPPPSWDMWKCLIRLKPRTLHG	1620
Qy	1621	PTPLLYRLGAVQNEVILTHPITIKYIMACMSADLEVVTSTWLVGVGLAALAAAYCLTTGVS	1680
Db	1621	PTPLLYRLGAVQNEVTTTHPITIKYIMACMSADLEVVTSTWLVGVGLAALAAAYCLTTGVS	1680
Qy	1681	VIVGRIILSGKPAVVPDRREVLYQBFDEMEBCASQLPYIEQGMQLAEQFKQKALGLLOTAT	1740
Db	1681	VIVGRIILSGKPAIIPDRREVLYREFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOTAT	1740
Qy	1741	KQAEAAAAPVSESKWRALETFWAKHWNFI SGIQYIAGLSTLPGNPAIASLMAFTASITSP	1800
Db	1741	KQAEAAAAPVSESKWRTLEAFWAKHWNFI SGIQYIAGLSTLPGNPAIASLMAFTASITSP	1800
Qy	1801	LTQNTLLFNILGGWAAQLAPPSSAASAFVAGAGTAGAAGVSGIGLKVILVDIILAGYGAGVA	1860
Db	1801	LTQNTLLFNILGGWAAQLAPPSSAASAFVAGAGTAGAAGVSGIGLKVILVDIILAGYGAGVA	1860
Qy	1861	GALVAFKVMSEVPSTEDLVNLLPALISPGALVGVWCAAIILRRHVGPGECAVQWNNRLI	1920
Db	1861	GALVAFKVMSEMPSTEDLVNLLPALISPGALVGVWCAAIILRRHVGPGECAVQWNNRLI	1920
Qy	1921	AFASRGNHVSPTHVYVPESDAAARVTQILSSITITOLLKRLHOWINEDCSTPCSGSWLRDV	1980
Db	1921	AFASRGNHVSPTHVYVPESDAAARVTQILSSITITOLLKRLHOWINEDCSTPCSGSWLRDV	1980
Qy	1981	WDWICTVLTDKFTWLOSLLIPLPCVPFLSCORGKGVWGRDGINOTTCCCAOITAGHVK	2040
Db	1981	WDWICTVLTDKFTWLOSLLIPLPCVPFLSCORGKGVWGRDGINOTTCCCAOITAGHVK	2040
Qy	2041	NGSMRI VGPRTCSNTWHGTFPIINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDPH	2100
Db	2041	NGSMRI VGPRTCSNTWHGTFPIINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDPH	2100
Qy	2101	YVTGMTDNVNCPCQVPAPEFTEVDGVR LHRYPACKPLIREDDVTFOGLNQYLVGSOL	2160
Db	2101	YVTGMTDNVNCPCQVPAPEFTEVDGVR LHRYPACKPLIREEVTVLGLNQYLVGSOL	2160
Qy	2161	PCPEPDDVLTLSMLTDP SHITAEAKERLARGSPPSLASSASQLSAPSLKATCTTHHD	2220
Db	2161	PCPEPDDVAVLTLSMLTDP SHITAEAKERLARGSPPSLASSASQLSAPSLKATCTTHHD	2220
Qy	2221	SPDADLIEANLLWROEMGNITRVSESNKWILDSFEPLHABGDEREISVAABEILKRSK	2280
Db	2221	SPDADLIEANLLWROEMGNITRVSESNKWILDSFEPLQABEEDEREVSVAEILRRSRK	2280
Qy	2281	FPSALPIWARPDYNPPLLESWKDPDYPVPHVGCPLPPTKAPPIPPPRKRRTVIVTESNV	2340
Db	2281	FPRAMPIWARPDYNPPLLESWKDPDYPVPHVGCPLPPAKAPPIPPPRKRRTVIVLSESTV	2340
Qy	2341	SSALAEATKTFGSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPLEGECDPL	2400
Db	2341	SSALAEATKTFGSGSSAVDSGTATASPDQSDGDKGSDVESYSSMPLEGECDPL	2400
Qy	2401	SDGSMSTVSEASESDVWCSSMSYTWTGALITPCAAEESKPLINPLNSLLRHHNMVYAIT	2460

```
Db 2401 SDGWSVTSEASEDDVCCSMYSYTWGALTIPCAABETKLPINALSNSLLRHHNLVYATT 2460
Qy 2461 SRASLRQKKVTFDRLOVLDHVDVLEKEMKASTVAKLLSIEBACKLTPPHSAKSKF 2520
Db 2461 SRASLRQKKVTFDRLOVLDHVDVLEKEMKASTVAKLLSIEBACKLTPPHSAKSKF 2520
Qy 2521 GYGAKOVNLSRAVNHRSWEDLLEDTPDTTIDTTIMAKSEVFCVQPEKGGRRKPARLIV 2580
Db 2521 GYGAKOVNLSRAVNHRSWEDLLEDTPDTTIDTTIMAKSEVFCVQPEKGGRRKPARLIV 2580
Qy 2581 FPDLGVRVCEKMAALYDVVSTLPQAVMGSSYGFQSPKQRVEFLVNTWKSCKKCPMGFSYDT 2640
Db 2581 FPDLGVRVCEKMAALYDVVSTLPQAVMGSSYGFQSPKQRVEFLVNTWKSCKKCPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVEESIYQCCDLAPARQAIRSLRTERLYIGPLTNSKQNCQYRCRA 2700
Db 2641 RCFDSTVTENDIRVEESIYQCCDLAPARQAIRSLRTERLYIGPLTNSKQNCQYRCRA 2700
Qy 2701 SGVLTTSCGNTLTCYLKATACRAKLODCTMLVNGDVLVVICESAGTQEDAAALRAFT 2760
Db 2701 SGVLTTSCGNTLTCYLKAAACRAKLODCTMLVCGDDLVVICESAGTQEDAAALRAFT 2760
Qy 2761 AMTRYSAPPDPPPEYDLELITSCSNVSVAHSDAGKRVVYLTRDPTPLARAAMETAR 2820
Db 2761 AMTRYSAPPDPPPEYDLELITSCSNVSVAHSDAGKRVVYLTRDPTPLARAAMETAR 2820
Qy 2821 HTPNSWLGNIIMYAPTLARMILMTHTFFSILLAOEKLADCOIYGACYSIEPLDLQP 2880
Db 2821 HTPNSWLGNIIMYAPTLARMILMTHTFFSILLAOEKLADCOIYGACYSIEPLDLQP 2880
Qy 2881 IIERLHGLSAPLTHSYSPGEINRVASCLKLGVPPLTRWHRARSVRAKLLSQGGRATC 2940
Db 2881 IIERLHGLSAPLTHSYSPGEINRVASCLKLGVPPLTRWHRARSVRAKLLSQGGRATC 2940
Qy 2941 GRYLFPNAVTRTKLTPTPAASQLDLSGWFVAGVSGGDIYHSLSRARPRWFPCLLLLSV 3000
Db 2941 GRYLFPNAVTRTKLTPTPAASQLDLSGWFVAGVSGGDIYHSLSRARPRWFPCLLLLSV 3000
Qy 3001 GVGIIYLLPNR 3010
Db 3001 GVGIIYLLPNR 3010

RESULT 13
ID ABG32460
AC ABG32460 standard; protein; 3010 AA.
XX
AC ABG32460;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate polyprotein mutant #9.
XX
KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 2199
FT /note= "Wild-type Ala substituted by Thr"
XX
XX WO200259321-A2.
XX
PD 01-AUG-2002.
XX
PF 16-JAN-2002; 2002WO-EP000526.
XX
PR 23-JAN-2001; 2001US-0263479P.
XX
PA (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
```

```
XX De Francesco R, Migliaccio G, Paonessa G;
PI WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX
PS Claim 1; Page; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polypeptide (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins), NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX
SQ Sequence 3010 AA;
```

```
Query Match 95.7%; Score 15319; DB 5; Length 3010;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 2851; Conservative 76; Mismatches 83; Indels 0; Gaps 0;

Qy 1 MSTNPKPQRTKNTNRRPQDVKEPGGQIVGGVYLLPRRGRPLGVATRKASRSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKEPGGQIVGGVYLLPRRGRPLGVATRKATRSRSQPRG 60

Qy 61 RRQPIPKARRPEGRWAQPGVPWPLYGNBGLGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRWAQPGVPWPLYGNBGLGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASAYEVNRVSGIYHVHTNDCSNSISVIEAADVIMHTPGCVPCVQEGNSSRCWV 240
Db 181 LLSCLTTPASAYEVNRVSGIYHVHTNDCSNSISVIEAADVIMHTPGCVPCVQEGNSSRCWV 240

Qy 241 ALPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYDGLCGSIFLVSQLFTEFSPRRHET 300
Db 241 ALPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYDGLCGSIFLVSQLFTEFSPRRHET 300

Qy 301 VQDCNCSIYPGHVSGHRMAMDMNWSPTTALVVSQLLRIPOAVVDMVAGAHGWGLAGLA 360
Db 301 VQDCNCSIYPGHVSGHRMAMDMNWSPTTALVVSQLLRIPOAVVDMVAGAHGWGLAGLA 360

Qy 361 YYSWVGNNWAKVLIIVALLFAGVDGTHTTGRVAGHTTSGFTSLFSSGASQKQLQVNTNGSW 420
Db 361 YYSWVGNNWAKVLIIVALLFAGVDGTHTTGRVAGHTTSGFTSLFSSGASQKQLQVNTNGSW 420

Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSCGPERMASCRIIDWFAQCWGGRITTYTKPNSS 480
Db 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSCGPERMASCRIIDWFAQCWGGRITTYTKPNSS 480
```

421	Db	HINRTALNCNDSLNTGFLAALPYVYHKFNSSGCPERMASCSPIDAPAQGWGPIITYNESHSS	480
481	Qy	DORPYCWHYAPPCGVPASQVCGPVYCTFSPVVGTTDRSGVPTYSWGENETDVMLLN	540
481	Db	DORPYCWHYAPPCGIVPAQVCGPVYCTFSPVVGTTDRFGVPTYSWGENETDVLN	540
541	Qy	NTRPQGNWFGCTWNNSTGFTKTCGGPPCNITGGVGNRTLICTDCFRKHPEATYTKCGSG	600
541	Db	NTRPPQGNWFGCTWNNSTGFTKTCGGPPCNIGGIGNKTLTCTDCFRKHPEATYTKCGSG	600
601	Qy	PWLTPRCLVDYRYLWHYPCNTLNFSPKVRMYGVGVEHRLNAACNWTFGERCNLEDRDRS	660
601	Db	PWLTPRCLVHYRYLWHYPCNTNFIFKVRMYGVGVEHRLNAACNWTFGERCNLEDRDRS	660
661	Qy	ELSPILLSTTEWOILPCAFPTLLPALSTGLIHLHQNIVDVQVLYGVGSFVFAIKWEYIL	720
661	Db	ELSPILLSTTEWOVLPSCFTLLPALSTGLIHLHQNVDVQVLYGIGSAVVSFAIKWEYVL	720
721	Qy	LLFLLADARVACACILMMMLLTAQEAALLENVLNNAASVAGAHGILSLFVFFCAAWYIKG	780
721	Db	LLFLLADARVACACILMMMLLTAQEAALLENVLNNAASVAGAHGILSLFVFFCAAWYIKG	780
781	Qy	RLAPGAAFYAGVWPELILLLLALPPRAVALDREMAASCGGAVLVGLVFLTLPSPYKVPLT	840
781	Db	RLVPGAAYALYGVWPELILLLLALPPRAYAMDREMAASCGGAVFVGLILLTLPSPHYKLFLA	840
841	Qy	RLIWLQVPIITRAEAHQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGLPLM	900
841	Db	RLIWLQVPIITRAEHLQVWVPLNVRGGRDAVILLTCAIHPELIFTITKILLAILGLPLM	900
901	Qy	VLQAGITRVPYEVRAQGLIRACMLVRKVAGGHYQVFMKLGALTGTVYVYNHLTPLRDWA	960
901	Db	VLQAGITKVPYFVRAHGLIRACMLVRKVAGGHYQVQMLMKLAALTGTVYVDHLTPLRDWA	960
961	Qy	HAGLRDLAVAPVPVFSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG	1020
961	Db	HAGLRDLAVAPVPVFSDMETKVIITWGADTAACGDIILGLPVSARRGREIHLGPADSLEG	1020
1021	Qy	QGWRLIAPITAYSQOTRGVLGCIITSLTRCDKNQVGEVQVVSATQSFATCINGVCWT	1080
1021	Db	QGWRLIAPITAYSQOTRGLLGCIITSLTRGRDNQVGEVQVVSATQSFATCINGVCWT	1080
1081	Qy	VYHGAGSKTLAGPKGPIITQMYTNVDLVLGVQAPPGARSMTPCSGSSDLXLVTRHADVI	1140
1081	Db	VYHGAGSKTLAGPKGPIITQMYTNVDQDLVGWQAPPGARSLTCTCGSSDLXLVTRHADVI	1140
1141	Qy	PVRRGDSRGSLLSRPVSYLKSGSGPILLCPSGHHVGVFRAAVCTRGVAKAVDPIPVES	1200
1141	Db	PVRRGDSRGSLLSRPVSYLKSGSGPILLCPSGHAGVIFRAAVCTRGVAKAVDPPVPVES	1200
1201	Qy	METTWRSVFTDNSTPPAVPOTFOVAHLHAPTSGSKSTKVPAAAYAAQGVKVLNPSVAA	1260
1201	Db	METTWRSVFTDNSTPPAVPOTFOVAHLHAPTSGSKSTKVPAAAYAAQGVKVLNPSVAA	1260
1261	Qy	TLGFGAYMSKANGIDPNIRGTVRTITGGSITYTYGKFLADGCGSGGAYDIILCDECHS	1320
1261	Db	TLGFGAYMSKANGIDPNIRGTVRTITGAPITYTYGKFLADGCGSGGAYDIILCDECHS	1320
1321	Qy	TDSTTILGIGTVDQAETAGARLVLAATPPGSVTVPHPNIEEIGLSNNGEIPFYGKAI	1380
1321	Db	TDSTTILGIGTVDQAETAGARLVLAATPPGSVTVPHPNIEEVALSSTGEIPFYGKAI	1380
1381	Qy	PTEAIKGRHLLIFCHSKKKDELAACLITGLIGNAVAYRGLDVSUIPPIGDVVVATDAL	1440
1381	Db	PTEIITKGRHLLIFCHSKKKDELAACLITGLIGNAVAYRGLDVSUIPPSGDVVVATDAL	1440
1441	Qy	MTGFTGDFDSVDCNTCVTQTVDFSLDPTFTTETTVPODAVRSQRGRTRGRGSGIYR	1500
1441	Db	MTGFTGDFDSVDCNTCVTQTVDFSLDPTFTTETTVPODAVRSQRGRTRGRGSGIYR	1500
1501	Qy	FVTPGERPSGMFDSVLCYCDAGCAWYELTPAETSVRLRAYLNTPGLPVCODHLEFPWES	1560
1501	Db	FVTPGERPSGMFDSVLCYCDAGCAWYELTPAETSVRLRAYLNTPGLPVCODHLEFPWES	1560
1561	Qy	VFTGLTHIDAHFLSQTQAGDNFPYLVAYQATVCARAQAAPPSPSWDQMWKCLIRLKLPTLHG	1620
1561	Db	VFTGLTHIDAHFLSQTQAGDNFPYLVAYQATVCARAQAAPPSPSWDQMWKCLIRLKLPTLHG	1620
1621	Qy	PTPELLYRLGAVQNEVILTHPIITKYIMACHMSADLEVVTSTWLVGGVLAALAAAYCLTGTGSV	1680
1621	Db	PTPELLYRLGAVQNEVITTHPIITKYIMACHMSADLEVVTSTWLVGGVLAALAAAYCLTGTGSV	1680
1681	Qy	VIVGRIILSGKPAVDPREVLYQEPDEMECEASQLPYIEQGMQLABQFKQKALGLLQATAT	1740
1681	Db	VIVGRIILSGKPAIIPREVLYREFDEMECEASHLPYIEQGMQLABQFKQKALGLLQATAT	1740
1741	Qy	KQAAAAAPVYESKWRALETFMAKHMWNFISGIOYLAGLSTLPCNPAIASLMAFTASITSP	1800
1741	Db	KQAAAAAPVYESKWRLEAFMAKHMWNFISGIOYLAGLSTLPCNPAIASLMAFTASITSP	1800
1801	Qy	LTTQNTLLFNILGGWAAQLAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVA	1860
1801	Db	LTTQHTLLFNILGGWAAQLAPPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVA	1860
1861	Qy	GALVAFKWSGVEPSTEDLVNLLPALTSPCALVGVVCAAILRRHRRHVGCEGAVQVMNRLI	1920
1861	Db	GALVAFKWSGEMPESTEDLVNLLPALTSPCALVGVVCAAILRRHRRHVGCEGAVQVMNRLI	1920
1921	Qy	AFASRGNHVSPTHYVPESDAAARVTOILSSLTITQLKRLHWINEDCSTPCSGSWMLRDV	1980
1921	Db	AFASRGNHVSPTHYVPESDAAARVTOILSSLTITQLKRLHWINEDCSTPCSGSWMLRDV	1980
1981	Qy	WDMICTVLDTFKTWLQSKLLPRLPGVPFFISCQRYGKVGWRGDGIMOTPCGAQIAGHVK	2040
1981	Db	WDMICTVLDTFKTWLQSKLLPRLPGVPFFISCQRYGKVGWRGDGIMOTPCGAQITGHVK	2040
2041	Qy	NGSMRIVGPTCNTWHGTTPINAYTTGCTPSPAPNYSRALWRVAAEVEVTRVCGDFH	2100
2041	Db	NGSMRIVGPTCNTWHGTTPINAYTTGCTPSPAPNYSRALWRVAAEVEVTRVCGDFH	2100
2101	Qy	YVGTGMTTDNVKPCQVPAPPEFFTEVDGVLHRVAPACKPLLRDDVTFQVGLNQLYLVGSOL	2160
2101	Db	YVGTGMTTDNVKPCQVPAPPEFFTEVDGVLHRVAPACKPLLRREVTFVGLNQLYLVGSOL	2160
2161	Qy	PCPEPEPDVTLTSMLTDPSHITAETAKRRLARGSPPLSSASSASQASAPSLKATCTTHHD	2220
2161	Db	PCPEPEPDVALTSMLTDPSHITAETAKRRLARGSPPLSTSSASASQASAPSLKATCTTRHD	2220
2221	Qy	SPADLILEANLLWRQEWGGINITRVESENKVILDSPEPLHAEGDEREISVAAEILARSRK	2280
2221	Db	SPADLILEANLLWRQEWGGINITRVESENKVILDSPEPLQAEEDEREVSVPAEILARSRK	2280
2281	Qy	PPSALPIWARPDYNNPILLESKWDPDVVPVHVHGCPLPPTKAPPIPPRRKRTTVVLTESNV	2340
2281			

```
Qy 2641 RCFDSTVTSIRVEESIQCCDLAPBARQAIRSLTERLYTGGPLTNSKGONCGYRRCRA 2700
Db 2641 RCFDSTVTENDIRVEESIQCCDLAPBARQAIRSLTERLYTGGPLTNSKGONCGYRRCRA 2700
Qy 2701 SGVLTTSCGNLTCLYKATACRAKLODCTMLVNGDDLVIICESAGTOEDAAALAPTE 2760
Db 2701 SGVLTTSCGNLTCLYKAAACRAKLODCTMLVCGDDLVIICESAGTOEDASLRAPTE 2760
Qy 2761 AMTRYAPPDPPOPEYDLELITSCSNVSVVAHDASGRVYVYLRDPTTTLARAAMETAR 2820
Db 2761 AMTRYAPPDPDPPEYDLELITSCSNVSVVAHDASGRVYVYLRDPTTTLARAAMETAR 2820
Qy 2821 HTPNSMIGNIIMVAPTLWARMILMTHFFSILLAQEOLKALDCQIYACYSIIEPLDLQ 2880
Db 2821 HTPVNSMIGNIIMVAPTLWARMILMTHFFSILLAQEOLKALDCQIYACYSIIEPLDLQ 2880
Qy 2881 IIERLHGLSARTLHYSRGEINRVASCLKLGVPPLRTWHRARSVRKLLSQGGRATC 2940
Db 2881 IIORLHGLSASFSLHYSRGEINRVASCLKLGVPPLRVWHRARSVRARLLSQGGRATC 2940
Qy 2941 GRYLPFNWAVRKLKLTPTPAASQDLDSGFVAGYSGGDIYHLSRARPWFPLCILLISV 3000
Db 2941 GRYLPFNWAVRKLKLTPTPAASQDLDSGFVAGYSGGDIYHLSRARPWFPMWCLLLISV 3000
Qy 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010
ID ABG32454 standard; protein; 3010 AA.
XX AC ABG32454;
XX AC
XX DT 15-NOV-2002 (first entry)
XX DE Hepatitis C virus Con 1 isolate polyprotein mutant #3.
XX KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;
XX KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1347 /note= "wild-type Ala substituted by Thr"
XX PN WO200259321-A2.
XX PN 01-AUG-2002.
XX PF 16-JAN-2002; 2002WO-EP000526.
XX PR 23-JAN-2001; 2001US-0263479P.
XX XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX PA De Francesco R, Migliaccio G, Paonessa G;
XX PI WPI; 2002-599793/64.
XX DR
XX PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX PT ribosome entry site (IRES) region, useful in studying HCV replication and
XX PT expression.
XX PS Claim 1; Page; 69pp; English.
XX XX The invention relates to nucleic acid molecules comprising altered HCV
XX CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
```

```
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC an HCV replicon enhanced cell or which containing a functional HCV replicon; (5)
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins), NS3 mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX SQ Sequence 3010 AA;
Query Match 95.7%; Score 15319; DB 5; Length 3010;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 2851; Conservative 76; Mismatches 83; Indels 0; Gaps 0;
Qy 1 MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRRGLGVRAKTRKASERSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRRGLGVRAKTRKTSERSQPRG 60
Qy 61 RRQIPKARPEGRAWAQPGYPWPPLYGNEGGLWAGLLSPRSPSPSGPTDPRRRSNLG 120
Db 61 RRQIPKARPEGRAWAQPGYPWPPLYGNEGGLWAGLLSPRSPSPSGPTDPRRRSNLG 120
Qy 121 KVDTLTCGFADLNGIYPLVGAAPLGAARALAHGVLEEDGVNATGNLPCGSFIFLLA 180
Db 121 KVDTLTCGFADLNGIYPLVGAAPLGAARALAHGVLEEDGVNATGNLPCGSFIFLLA 180
Qy 181 LLSCLTTPASAYEVNRVSGIYHVTNDCSNSSIVYEAADVIHMTPCVPCVCGNSSRCWV 240
Db 181 LLSCLTTPASAYEVNRVSGIYHVTNDCSNSSIVYEAADVIHMTPCVPCVCGNSSRCWV 240
Qy 241 ALPTLAARNASVPTTIRRHVDLLVGTAAFCFSAMYVGLCGSIFLVSQLFSPRRHET 300
Db 241 ALPTLAARNASVPTTIRRHVDLLVGTAAFCFSAMYVGLCGSIFLVSQLFSPRRHET 300
Qy 301 VQDCNCSIYPGHVSGHRMAMDMNWSPTTALVVSQLLRIPQAVVDMVAGAHWGLAGLA 360
Db 301 VQDCNCSIYPGHVSGHRMAMDMNWSPTTALVVSQLLRIPQAVVDMVAGAHWGLAGLA 360
Qy 361 YYSWGNWAKULIVALLFAGVDGTHVTGRVAGHTTSGTSLFSSGASOKQLVNTNGSW 420
Db 361 YYSWGNWAKULIVALLFAGVDGTHVTGRVAGHTTSGTSLFSSGASOKQLVNTNGSW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSCCPERMASCRIPIIDFAQGWGITTKPNSS 480
Db 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSCCPERMASCRIPIIDFAQGWGITTKPNSS 480
Qy 481 DQRPYCHVYAPRCGVVPASQVCFVYCTPSPVVVGTGTTDRSGVPTYSWGENETDVMLLN 540
Db 481 DQRPYCHVYAPRCGVVPASQVCFVYCTPSPVVVGTGTTDRSGVPTYSWGENETDVMLLN 540
Qy 541 NTRPQGNWFGCTWMNSTGFTKTCGGPPCNTGGVGNRTLICPTCFRKHPEATYTKCSG 600
Db 541 NTRPQGNWFGCTWMNSTGFTKTCGGPPCNTGGVGNRTLICPTCFRKHPEATYTKCSG 600
Qy 601 PWLTFRCLVDYPRYLWHYPCTLNFSIFKVRMYVGVGVEHRLNAACNWTGERCNLEDRRS 660
Db 601 PWLTFRCLVDYPRYLWHYPCTLNFSIFKVRMYVGVGVEHRLNAACNWTGERCNLEDRRS 660
```

QY 661 ELSPLLSTTEWQIILPCAFTTILPALSTGLIHJHQNIVDVQYLYGVGSAPVSPFAIKWEYIL 720
Db 661 ELSPLLSTTEWQVLPSCFTTILPALSTGLIHJHQNVDVQYLYGIGCVSVSPFAIKWEYVL 720
QY 721 LFLFLLLADARVCACILWMLLIIAQEAALENIVLNAASVAGAHGILSLFVFCFAAWYIKG 780
Db 721 LFLFLLLADARVCACILWMLLIIAQEAALENIVLNAASVAGAHGILSLFVFCFAAWYIKG 780
QY 781 RLAPGAAYAFYGVWPLLLLLLLPRAVALDREMAASCGLVGLVFLTILSPYKVFIL 840
Db 781 RLVPGAAYALXGWPMLLLLLLLPRAVAMDREMAASCGLVFGVGLILLLTSPHYKFLA 840
QY 841 RLILWLQYFITRAEAHQWVPLNVRGGRDAIILLTCAVHPBELIPDITKLLAILGLPILM 900
Db 841 RLILWLQYFITRAEAHLQWIPPLNVRGGRDAVILLTCAIHPELIPTTIKILLAILGLPILM 900
QY 901 VLOAGITRVPYFVRAOGLIRACMLVRKVAGHYQVFMKLCALTCVYVYHLLTPLRDWA 960
Db 901 VLOAGITRVPYFVRAHGLIRACMLVRKVAGHYQVQALMKLAALTGTVYVYDHLTPLRDWA 960
QY 961 HAGLRDLAVAVPVPVSAMETKVIITWGADTAACGDDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HAGLRDLAVAVPVPVSDMETKVIITWGADTAACGDDIILGLPVSARRGHEIHLGPADSLEG 1020
QY 1021 QOWRLAPITAYSQOTRGVLGCIITSLTGRDKNQVEGEVQVYSTATQSFPLATCINGVCWT 1080
Db 1021 QOWRLAPITAYSQOTRGLGCCIITSLTGRDRNQVEGEVQVYSTATQSFPLATCVNGVCWT 1080
QY 1081 VYHGAGSKTLAPKGPITOMYTNVDLILVWQAPPGARSMTPCSGSSDLYLVYTRHADVI 1140
Db 1081 VYHGAGSKTLAPKGPITOMYTNVDQLVWQAPPGARSLTPTCTCGSSDLYLVYTRHADVI 1140
QY 1141 PYRRRGDSRGSLLSPRPVSYLKGSSGGLLPCSGHVGVFRAAVCTRGVAKAVDPIPVES 1200
Db 1141 PYRRRGDSRGSLLSPRPVSYLKGSSGGLLPCSGHAGVIFRAAVCTRGVAKAVDVPVDES 1200
QY 1201 METTMRSPVFTDNSTPPAVPOTFOVAHLHAPTGSKGSKTKVPAAYAAQGYKVILVLPNSVAA 1260
Db 1201 METTMRSPVFTDNSSPPAVPOTFOVAHLHAPTGSKGSKTKVPAAYAAQGYKVILVLPNSVAA 1260
QY 1261 TLGFGAYMSKAHGDINIRGTVRTITGSGIITYSTYKFLADGGCSGGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGDINIRGTVRTITGAPITYSTYKFLADGGCSGGAYDIIICDECHS 1320
QY 1321 TDSITILGTVLDQETAGARLVLATATPGSVTVPHPNIEETGLSNNGEIPFVGKAI 1380
Db 1321 TDSITILGTVLDQETAGARLVLTATPGSVTVPHPNIEEVALSSTGEIPFVGKAI 1380
QY 1381 PIEAIKGRHLLIFCHSKKKCDELAALKLGLGLNAVAYYRGLDVSVIPTGDDVVVATDAL 1440
Db 1381 PIETIKGRHLLIFCHSKKKCDELAALKLGLGLNAVAYYRGLDVSVIPTSGDIVVATDAL 1440
QY 1441 MTGFTGDFOSVDCNTCVTQTVDFSLDPTFTTETTVPODAVSRSGRRGRTGRSGIYR 1500
Db 1441 MTGFTGDFOSVDCNTCVTQTVDFSLDPTFTTETTVPODAVSRSGRRGRTGRGMGIYR 1500
QY 1501 FVTPGERPSGMPDSSVLCBCEYDAGCAWYELTPAETSVRILRAYLNTPLGVPVQDHLFEWES 1560
Db 1501 FVTPGERPSGMPDSSVLCBCEYDAGCAWYELTPAETSVRILRAYLNTPLGVPVQDHLFEWES 1560
QY 1561 VFTGLTHIDAHFLSOTKQAGDNPPYLVAQVATVCARAQAPPPSWDMWKCLRKLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSOTKQAGDNPPYLVAQVATVCARAQAPPPSWDMWKCLRKLKPTLHG 1620
QY 1621 PTPPLYRLGAVQNEVILTHPIITKYIMACMSADLEVVTSTVWLIVGGVLAALAAAYCLTITGVS 1680
Db 1621 PTPPLYRLGAVQNEVITTHPIITKYIMACMSADLEVVTSTVWLIVGGVLAALAAAYCLTITGVS 1680
QY 1681 VLVGRILISGKPAVDPDRVILYQEFDEMEECASQLPYIEQGMQLABQFKOKALGILLOTAT 1740
Db 1681 VLVGRILISGKPAIIPDRVILYREFDEMEECASHLPYIEQGMQLABQFKOKAIGLLOTAT 1740
QY 1741 KOAEAAAAPVVEKSWRALETFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800

Db 1741 KOAEAAAAPVVEKSWRTLEAFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
QY 1801 LTTQNTLLFNILGGWVAQAAPPAAASAFVAGIAGAAGVSGIGLKVLVDILAGYAGVA 1860
Db 1801 LTTQNTLLFNILGGWVAQAAPPAAASAFVAGIAGAAGVSGIGLKVLVDILAGYAGVA 1860
QY 1861 GALVAFKVMSEVPSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPEGAVQWNNRLI 1920
Db 1861 GALVAFKVMSEMPSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPEGAVQWNNRLI 1920
QY 1921 APASRGNHVSPTHVPESDAAARVTOILSSITITOLLKRLHOWINEDCSTPCSGSWLRDV 1980
Db 1921 APASRGNHVSPTHVPESDAAARVTOILSSITITOLLKRLHOWINEDCSTPCSGSWLRDV 1980
QY 1981 WDWICTVLTDFTKTLQSKLLPRLPGVPFLSCQGYKGVWRGGINQTTTCCPQAOIAGHVK 2040
Db 1981 WDWICTVLTDFTKTLQSKLLPRLPGVPFLSCQGYKGVWRGGINQTTTCCPQAOIAGHVK 2040
QY 2041 NGSMRIVGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDPH 2100
Db 2041 NGSMRIVGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDPH 2100
QY 2101 YVTGMTTNDVKPCQVPAPEPTEVDGVRILHRYAPACKPLLBEDVTFOVGLNQYLVGSOL 2160
Db 2101 YVTGMTTNDVKPCQVPAPEPTEVDGVRILHRYAPACKPLLBEEVTVLVLNQYLVGSOL 2160
QY 2161 PCEPEPDTVLTLSMLTDPSSHITTAETAKRRLARGSPPSLASSASOLSAPSLKATCTTHHD 2220
Db 2161 PCEPEPDTVAVLUTSMLTDPSSHITTAETAKRRLARGSPPSLASSASOLSAPSLKATCTTHHD 2220
QY 2221 SPDADLIEANLLWRQEMGNITRVESENKVILDSFEPLOAEEDEREVSVPAEILRRSRK 2280
Db 2221 SPDADLIEANLLWRQEMGNITRVESENKVILDSFEPLOAEEDEREVSVPAEILRRSRK 2280
QY 2281 FPSALPIWARDYNPPLLESWKDPDYVPVHVGCPLPPTKAPPIPPRRKRTVVLTESV 2340
Db 2281 FPRAMPPIWARDYNPPLLESWKDPDYVPVHVGCPLPPAKAPPIPPRRKRTVVLSESTV 2340
QY 2341 SSALAEALATKTFGSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPPLGEQDPL 2400
Db 2341 SSALAEALATKTFGSESSAVDSGTATASDPDSDGDKGSDVESYSSMPPLGEQDPL 2400
QY 2401 SDGSMSTVSEASESDVVCSSSYWTGALITPCAASESKLPINPLSNSLLRHHNMVYAT 2460
Db 2401 SDGSMSTVSEASESDVVCSSSYWTGALITPCAABETKLPINALSNSLLRHHNLVYAT 2460
QY 2461 SRSASLRQKVTFFDLQVLDHHRDVLKEMKAKASTVKAKLLSIBEAACKLTPPHSAKSF 2520
Db 2461 SRSASLRQKVTFFDLQVLDHHRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF 2520
QY 2521 GYGAKDVNLSRAVNHTRSVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGKPARLIV 2580
Db 2521 GYGAKDVNLSKSAVNHTRSVWMDLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIV 2580
QY 2581 PFDLGVRCXKMALDVVSTLPOAVMGSSYGFQYSPKORVEPLVNTWKSCKCPMFPST 2640
Db 2581 PFDLGVRCXKMALDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWKAKKCPMGFAYT 2640
QY 2641 RCFDSTVTESDIRVEESYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRCA 2700
Db 2641 RCFDSTVTENDIRVEESYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRCA 2700
QY 2701 SGVLTSTCGNLTLYLKATAACRAAKLODCTMLVNGDDLVTICESAGTOEDAAALRAFTE 2760
Db 2701 SGVLTSTCGNLTLYLKAAAACRAAKLODCTMLVCGDDLVTICESAGTOEDASARAFTE 2760
QY 2761 AMTRYSAAPGPPPOPEYDLELITSCSSNVSVVAHDASGKRVYVLTDPPTPLARAAWETAR 2820
Db 2761 AMTRYSAAPGPPPKPEYDLELITSCSSNVSVVAHDASGKRVYVLTDPPTPLARAAWETAR 2820
QY 2821 HTTPINSWLGNITIMYAPTILWARMILMTHFPFILLSAQOLEKALDCQIYGACYSIBPLDLPO 2880

Db 2821 HTPVNSWLGNIIMVAPTLWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQ 2880
Qy 2881 IIERLHGLSFTLHSGSPGGINRVASCLRLGVPPLKRTWHRARSVRKLLSOGGRAATC 2940
Db 2881 IIOQLHGLSFTLHSGSPGGINRVASCLRLGVPPLKRTWHRARSVRKLLSOGGRAATC 2940
Qy 2941 GRYLFNMAVTRKLTLPAPASQDLSGFWFVAGSGGDIYHSLSRAPRPWPICLLLLSV 3000
Db 2941 GRYLFNMAVTRKLTLPAPASQDLSGFWFVAGSGGDIYHSLSRAPRPWPICLLLLSV 3000
Qy 3001 GVGIIYLLPNR 3010
Db 3001 GVGIIYLLPNR 3010

RESULT 15

ABG32461
ID ABG32461 standard; protein; 3010 AA.
XX AC ABG32461;
XX 15-NOV-2002 (first entry)
XX Hepatitis C virus Con 1 isolate polyprotein mutant #10.
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NSSA; HCV replication; polyprotein.
XX Hepatitis C virus.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 2204 /note= "Wild-type Ser substituted by Arg"
PT
XX WO200259321-A2.
XX 01-AUG-2002.
XX 16-JAN-2002; 2002WO-EP000526.
XX 23-JAN-2001; 2001US-0263479P.
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX Claim 1; Page; 69pp; English.
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NSSA, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound

CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins), NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX
SQ Sequence 3010 AA;
Query Match 95.7%; Score 15318; DB 5; Length 3010;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 2851; Conservative 76; Mismatches 83; Indels 0; Gaps 0;
Qy 1 MSTNPKPQRTKNTNRRPQDVKFGGGQIVGGVYLLPRRGLPRGVRATKASERSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKFGGGQIVGGVYLLPRRGLPRGVRATKASERSQPRG 60
Qy 61 RRQPIPKARRPEGRAWAQPGYPMPLYGNEGLWAGWLLSPRGRSPGWPTDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRAWAQPGYPMPLYGNEGLWAGWLLSPRGRSPGWPTDPRRRSRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNATGNTLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNATGNTLPGCSFSIFLLA 180
Qy 181 LLSCLTIPASAYERNVSGIYHVTNDCNSSIYVEAADVTMTGTCVPCVQEGNSSRCWV 240
Db 181 LLSCLTIPASAYERNVSGIYHVTNDCNSSIYVEAADVTMTGTCVPCVQEGNSSRCWV 240
Qy 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLGGSGIFLVSQLFTSPRRHET 300
Db 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLGGSGIFLVSQLFTSPRRHET 300
Qy 301 VQDCNCSIYPGHVSCHRMAMDMNMNSPTTALVVSQLLRIPQAVDMVAGAHGVLGAGLA 360
Db 301 VQDCNCSIYPGHVSCHRMAMDMNMNSPTTALVVSQLLRIPQAVDMVAGAHGVLGAGLA 360
Qy 361 YYSNVGNWAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW 420
Db 361 YYSNVGNWAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCCPERMASCPIPWFAQGWGPITTKNSS 480
Db 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCCPERMASCPIPWFAQGWGPITTKNSS 480
Qy 481 DORPYCHYAPRCGVVPASQVCPVYCTPSPVVVGTTRSDRSGVPTYSWGENETDVLNLN 540
Db 481 DORPYCHYAPRCGVVPASQVCPVYCTPSPVVVGTTRSDRSGVPTYSWGENETDVLNLN 540
Qy 541 NTRPQGNWFGCTWMNSTGFTKTCGPPCNIGGVGNRTLCPTDFRKHPEATYTKCGSG 600
Db 541 NTRPQGNWFGCTWMNSTGFTKTCGPPCNIGGVGNRTLCPTDFRKHPEATYTKCGSG 600
Qy 601 PWLTPRCCLVDYPRYLMHYPCITLNFSEIFKVRMYGVGVHRLNAACNWTGERCNLEDRDRS 660
Db 601 PWLTPRCCLVDYPRYLMHYPCITLNFSEIFKVRMYGVGVHRLNAACNWTGERCNLEDRDRS 660
Qy 661 ELSPLLLSTTEWQLPCAFITLTPALSTGLIHLHQNIVDVQVLYGVGSAFVFAIKWEYIL 720
Db 661 ELSPLLLSTTEWQLPCAFITLTPALSTGLIHLHQNIVDVQVLYGVGSAFVFAIKWEYIL 720
Qy 721 LFLFLLADARVACILWMLLIAQBALENIWLNAAASVAGAHGILSFLVFFCAAWYIKG 780
Db 721 LFLFLLADARVACILWMLLIAQBALENIWLNAAASVAGAHGILSFLVFFCAAWYIKG 780
Qy 781 RLAPGAAYFVGVWPLLLHLLALPPRAYALDRENAASCGGAVLGLVFLTSPYKVLFT 840
Db 781 RLAPGAAYFVGVWPLLLHLLALPPRAYALDRENAASCGGAVLGLVFLTSPYKVLFT 840
Qy 841 RLWMLQVFIYRAEAMQVWVPPNVNRGRDAIILLTCAVHPELITFITKLLAILGLPLM 900
Db 841 RLWMLQVFIYRAEAMQVWVPPNVNRGRDAIILLTCAVHPELITFITKLLAILGLPLM 900

Db 841 RLIWMLQYFITRAEHLQVWIPLPNVRGGRDAVILLTCAIHPELIFTITKILLAILGLPLM 900
Qy 901 VLQAGITRVPYFVRAQGLIRACMLVRKVAGGHYQVFMKLGALGTGYVYNHLTPLRDWA 960
Db 901 VLQAGITKVPYFVRAHGLIRACMLVRKVAGGHYQVFMKLAALGTGYVDHLTPLRDWA 960
Qy 961 HAGLRDLAVAPVVPVFSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HAGLRDLAVAPVVPVFSMETKVIITWGADTAACGDIILGLPVSARRGREIHILGPADSLEG 1020
Qy 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDNQVEGEVQVYSTATQSFATCINGVCWT 1080
Db 1021 QGWRLLAPITAYSQOTRGLLGCIIISLGRDRNQVEGEVQVYSTATQSFATCVNGVCWT 1080
Qy 1081 VYHGAGSKTLAGKGPITQMYTNVDLDLVGWQAPPGARSMTPCSGSDLYLVTRHADVI 1140
Db 1081 VYHGAGSKTLAGKGPITQMYTNVDQDLVGWQAPPGARSLTPTCTCGSDLYLVTRHADVI 1140
Qy 1141 PYRRGDSRGLLSRPPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDEIPVES 1200
Db 1141 PYRRGDSRGLLSRPPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVES 1200
Qy 1201 METTMRSPVFTDNSTPPAVPQTFQVAHLHAPTGSKGSTKVPAAQAQGYKVLVLNPSVAA 1260
Db 1201 METTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSKGSTKVPAAQAQGYKVLVLNPSVAA 1260
Qy 1261 TLGFGAYMSKAHGIDNIRTGVRTITGGSITYSTVGKFLADGGCGGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGIDNIRTGVRTITTGAPITYSTVGKFLADGGCGGAYDIIICDECHS 1320
Qy 1321 TDSSTTLTGTLVDQAETHAGRLVLATATPPGSVTVPHNPBIEBGLSNGBIPFYGKAI 1380
Db 1321 TDSSTTLTGTLVDQAETHAGRLVLATATPPGSVTVPHNPBIEEVALSSTGSEIPFYGKAI 1380
Qy 1381 PIETIKGGRHLIFCHSKKKCDELAALKTGLGNVAVYRGLOVSVITPGDVVVVATDAL 1440
Db 1381 PIETIKGGRHLIFCHSKKKCDELAALKTGLGNVAVYRGLOVSVITPGDVVVVATDAL 1440
Qy 1441 MTGFTGDFDSVDCNCTVQTFVDFSLDPTFTTETTVPODVSRSORRQRTGRGSGIYR 1500
Db 1441 MTGFTGDFDSVDCNCTVQTFVDFSLDPTFTTETTVPODVSRSORRQRTGRGSGIYR 1500
Qy 1501 FVTPGERPSGMFDSVLCBECYDAGCAWYELTPAETSVRLRAYLNTPLPVCQDHLFEWES 1560
Db 1501 FVTPGERPSGMFDSVLCBECYDAGCAWYELTPAETSVRLRAYLNTPLPVCQDHLFEWES 1560
Qy 1561 VFTGLTHIDAHFLSQTQKQAGDNFPYLVAQVATCARAQAPPPSDQMWKCLIRLXPTLHG 1620
Db 1561 VFTGLTHIDAHFLSQTQKQAGDNFPYLVAQVATCARAQAPPPSDQMWKCLIRLXPTLHG 1620
Qy 1621 PTPLLYRLGAVONEVLTTHPI TKYI WACMSADLEVVTSTWVLVGGVLAALAAAYCLTTSV 1680
Db 1621 PTPLLYRLGAVONEVTTTHPI TKYI WACMSADLEVVTSTWVLVGGVLAALAAAYCLTTSV 1680
Qy 1681 VIVGRILISGKPAVVPDREVLVYQEFDEMEBCASQLPIEQGMQLABQPFQKQALGLIQTAT 1740
Db 1681 VIVGRILISGKPAIIFDREVLVREFDEMEBCASHLPYIEQGMQLABQPFQKQALGLIQTAT 1740
Qy 1741 KQAEAAAPVVPESKWRALFTFWAKHMWNFTSGIQYLAGLSTLPCNPAIASIMAFATASITSP 1800
Db 1741 KQAEAAAPVVPESKWRLEAFWAKHMWNFTSGIQYLAGLSTLPCNPAIASIMAFATASITSP 1800
Qy 1801 LTTQNTLLFNILGGWAAQOLAPPSAASAFVGAGIAGAAGVSGICLGVLDVILLAGYGAGVA 1860
Db 1801 LTTQHTLLFNILGGWAAQOLAPPSAASAFVGAGIAGAAGVSGICLGVLDVILLAGYGAGVA 1860
Qy 1861 GALVAFKWSGEPSTEDLVNLLPAILLSPGALVGVVCAAILRRHVGPGEVQWNNRLI 1920
Db 1861 GALVAFKWSGEPSTEDLVNLLPAILLSPGALVGVVCAAILRRHVGPGEVQWNNRLI 1920
Qy 1921 AFASRGNHVSPTHYVPESDAAARVTOILSSLTITQLLKRHLQWINEBDCSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHYVPESDAAARVTOILSSLTITQLLKRHLQWINEBDCSTPCSGSWLRDV 1980

Qy 1981 WDWICTVLTDFTKTLQSKLLPRLPGVPFLSCORGYKGVNRRGDMOTTCPCGAQIAGHVK 2040
Db 1981 WDWICTVLTDFTKTLQSKLLPRLPGVPFLSCORGYKGVNRRGDMOTTCPCGAQITGHVK 2040
Qy 2041 NGSMRIVGPRCNSNTHGTFPINAATTGCTPTSPAPNYSRALWRVAAEYEVETVRGDFH 2100
Db 2041 NGSMRIVGPRCNSNTHGTFPINAATTGCTPTSPAPNYSRALWRVAAEYEVETVRGDFH 2100
Qy 2101 YVTGMTDNVCKPCQVPAPEFTEVDGVRHLHRYAPACKPLREDVTVFQVGLNQYLVSQOL 2160
Db 2101 YVTGMTDNVCKPCQVPAPEFTEVDGVRHLHRYAPACKPLREEVTVFVGLNQYLVSQOL 2160
Qy 2161 PCEPEPDPVTVTSMLTDPSSHITAETAKRRLARGSPPSLASSASQLSAPSLKATCTTHHD 2220
Db 2161 PCEPEPDPVAVLTSMLTDPSSHITAETAKRRLARGSPPSLASSARQLSAPSLKATCTTHRD 2220
Qy 2221 SPDADLIEANLLMRQEMGNITRVESENKVVILDSFEPLQAEEDREVSVPABEILURSRK 2280
Db 2221 SPDADLIEANLLMRQEMGNITRVESENKVVILDSFEPLQAEEDREVSVPABEILURSRK 2280
Qy 2281 FPSALPIWARPDYNPPILLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRRTVILTESNV 2340
Db 2281 FPRAMP IWARPDYNPPILLESWKDPDYVPVHVHGCPLPPAKAPPIPPRRKRRTVILSESV 2340
Qy 2341 SSALAEALATKTFGSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPPLEGEFGDPDL 2400
Db 2341 SSALAEALATKTFGSGSSAVDSGTATASDPQSDDDGAGSDVESYSSMPPLEGEFGDPDL 2400
Qy 2401 SDGWSVTSEASEBDDVCCSMSYTWGALITPCAARESKLPTNPLNSLLRHHNMVYATT 2460
Db 2401 SDGWSVTSEASEBDDVCCSMSYTWGALITPCAARESKLPTNPLNSLLRHHNMVYATT 2460
Qy 2461 SRSASLRQKVTFFORLOVLDHRYDVLKEMAKASTVAKALLSIBEACKLTPPHSAKSKF 2520
Db 2461 SRSASLRQKVTFFORLOVLDHRYDVLKEMAKASTVAKALLSIBEACKLTPPHSAKSKF 2520
Qy 2521 GYGAKDVNLSKSSAVNHIRSVWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVNLSKSSAVNHIRSVWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Qy 2581 FPDGLVRVCERKALYDVVSTLPOAVMGSSYGFQVSPKORVEFLVNTWKSXKCPMGFSDYT 2640
Db 2581 FPDGLVRVCERKALYDVVSTLPOAVMGSSYGFQVSPKORVEFLVNTWKSXKCPMGFSDYT 2640
Qy 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGNGCYRRCKRA 2700
Db 2641 RCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGNGCYRRCKRA 2700
Qy 2701 SGVLTTSCTGNTLTCVYLKATAACRAKLODCTMLVNGDDDLVWICESAGTQEDAAALRAFTE 2760
Db 2701 SGVLTTSCTGNTLTCVYLKATAACRAKLODCTMLVNGDDDLVWICESAGTQEDAAALRAFTE 2760
Qy 2761 AMTRYSAAPGPPPOPEYDLELITSCSSNVSAHADSGKRVYVLTDRDPTPLARAAWETAR 2820
Db 2761 AMTRYSAAPGPPPOPEYDLELITSCSSNVSAHADSGKRVYVLTDRDPTPLARAAWETAR 2820
Qy 2821 HTPINSWLGNTIMYAPTILWARMILMTHFSSILLAQEOLEKALDQOYGCACYSIEPLDLPO 2880
Db 2821 HTPINSWLGNTIMYAPTILWARMILMTHFSSILLAQEOLEKALDQOYGCACYSIEPLDLPO 2880
Qy 2881 IIERUHLGSAFTLHSGYSPGEINRVASCLURKLGVPPLRTWRHRARSVRKLLSQGGRAATC 2940
Db 2881 IIORUHLGSAFTLHSGYSPGEINRVASCLURKLGVPPLRVWRHRARSVRKLLSQGGRAATC 2940
Qy 2941 GRYLFNWAVRTKILKLTIPAAASQDLDSGFVAGYSGGDIYHLSLRARPRWFFPLCLLLLSV 3000
Db 2941 GRYLFNWAVRTKILKLTIPAAASQDLDSGFVAGYSGGDIYHLSLRARPRWFFPLCLLLLSV 3000
Qy 3001 GVGIIYLLPNR 3010
Db 3001 GVGIIYLLPNR 3010

Search completed: December 3, 2005, 08:32:50
Job time : 331 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	16009	100.0	3010	2	US-09-014-416-3	Sequence 3, Appli
2	15319	95.7	3010	2	US-09-539-601-3	Sequence 3, Appli
3	15289	95.5	3010	2	US-09-539-601-27	Sequence 27, Appl
4	15279	95.4	3010	2	US-09-539-601-21	Sequence 21, Appl
5	15269	95.4	3010	2	US-09-539-601-33	Sequence 33, Appl
6	15253	95.3	3010	1	US-08-324-977-2	Sequence 2, Appli
7	15253	95.3	3010	1	US-08-324-977-14	Sequence 14, Appl
8	15253	95.3	3010	1	US-08-384-616-2	Sequence 2, Appli
9	15253	95.3	3010	1	US-08-384-616-14	Sequence 14, Appl
10	15253	95.3	3010	1	US-08-304-686A-2	Sequence 2, Appli
11	15253	95.3	3010	1	US-08-304-686A-14	Sequence 14, Appl
12	15253	95.3	3010	2	US-09-315-850-2	Sequence 2, Appli
13	15253	95.3	3010	2	US-09-315-850-14	Sequence 14, Appl
14	14801.5	92.5	2985	2	US-10-259-275-40	Sequence 40, Appl
15	14640.5	91.5	3011	2	US-09-014-416-5	Sequence 5, Appli
16	14143.5	88.3	3012	2	US-08-811-566-2	Sequence 2, Appli
17	14143.5	88.3	3012	2	US-09-034-756-2	Sequence 2, Appli
18	14130.5	88.3	3011	2	US-08-850-328-1	Sequence 1, Appli
19	14124.5	88.2	3011	2	US-09-014-416-1	Sequence 1, Appli
20	14124.5	88.2	3011	2	US-09-352-572-9	Sequence 9, Appli
21	14124.5	88.2	3011	2	US-10-259-275-20	Sequence 20, Appl
22	14119.5	88.2	3011	1	US-08-440-103-36	Sequence 36, Appl
23	14119.5	88.2	3011	1	US-08-440-542-36	Sequence 36, Appl
24	14119.5	88.2	3011	1	US-07-910-760-10	Sequence 10, Appl
25	14119.5	88.2	3011	1	US-08-440-519-10	Sequence 10, Appl
26	14119.5	88.2	3011	1	US-08-231-368-36	Sequence 36, Appl
27	14119.5	88.2	3011	1	US-08-440-210-36	Sequence 36, Appl

Db 301 VODNCNSIYPGHVSHRMAWDMNMNWSPTTALVVSQLLRIPOAVVDMVAGAHWGVLAGLA 360
Qy 361 YYSVMGNWAKVLIIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSGSASOKIQLVNTNGSW 420
Db 361 YYSVMGNWAKVLIIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSGSASOKIQLVNTNGSW 420
Qy 421 HINRTALCNSDSLOTGFFAALFYAHKFNSSGCPERMASCRPIDWFAQWGPIITYTKPNSS 480
Db 421 HINRTALCNSDSLOTGFFAALFYAHKFNSSGCPERMASCRPIDWFAQWGPIITYTKPNSS 480
Qy 481 DORPYCWHYARPCGVVPASQVCGVYCFTPSPVVVGTDRSGVPTYSWGENETDVMLLN 540
Db 481 DORPYCWHYARPCGVVPASQVCGVYCFTPSPVVVGTDRSGVPTYSWGENETDVMLLN 540
Qy 541 NTRPPQGNWFGCTWNNSGTFTKCGPPCNTGGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Db 541 NTRPPQGNWFGCTWNNSGTFTKCGPPCNTGGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Qy 601 PWLTPRCLVDIYPYRLWHYPCTLNFSIFKVRMYVGVGEHRLNAAACNWTGERCNLEDRDRS 660
Db 601 PWLTPRCLVDIYPYRLWHYPCTLNFSIFKVRMYVGVGEHRLNAAACNWTGERCNLEDRDRS 660
Qy 661 ELSPLLLSTTEWQILPCAFITLIPALSTGLIHLHONIVDVQVLYGVGSAFVGFALIKWEYIL 720
Db 661 ELSPLLLSTTEWQILPCAFITLIPALSTGLIHLHONIVDVQVLYGVGSAFVGFALIKWEYIL 720
Qy 721 LFLLLADARVACALWMLLLIAQAEAALENIVLNAASVAGAHGILSLFVFFCAAWYIKG 780
Db 721 LFLLLADARVACALWMLLLIAQAEAALENIVLNAASVAGAHGILSLFVFFCAAWYIKG 780
Qy 781 RLAPGAAVAFYGVMPLELLLLALLPGRAYALDREMAASCGLVAVLGLVFLTLSPYYKVFILT 840
Db 781 RLAPGAAVAFYGVMPLELLLLALLPGRAYALDREMAASCGLVAVLGLVFLTLSPYYKVFILT 840
Qy 841 RLILWLQVFIITRAEAHMQVPPPLNVNRGGRDAIILLTCAVHPELIFDITKULLAILGLPLM 900
Db 841 RLILWLQVFIITRAEAHMQVPPPLNVNRGGRDAIILLTCAVHPELIFDITKULLAILGLPLM 900
Qy 901 VLQAGITRVPVFVRAOGLIRACMLVRKVAGGHYQVMFEMKLGALTGTVVYNHLPFLRDWA 960
Db 901 VLQAGITRVPVFVRAOGLIRACMLVRKVAGGHYQVMFEMKLGALTGTVVYNHLPFLRDWA 960
Qy 961 HAGLRDLAVAVPVVFSAMETKVITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HAGLRDLAVAVPVVFSAMETKVITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Qy 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDNQVGEVQVVSSTATQSFATCINGVCWT 1080
Db 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDNQVGEVQVVSSTATQSFATCINGVCWT 1080
Qy 1081 VYHGAGSKTLGAPKGPITOMVTNVDLDLVGHOAPPGARSMTPCSCGSSDLVLTVRHADVI 1140
Db 1081 VYHGAGSKTLGAPKGPITOMVTNVDLDLVGHOAPPGARSMTPCSCGSSDLVLTVRHADVI 1140
Qy 1141 PVRRRGDSRGSLLSPRPVSYLKGSGGGLLPCPSGHVGVFVRAAIVCTRCVAKAVDFIPVES 1200
Db 1141 PVRRRGDSRGSLLSPRPVSYLKGSGGGLLPCPSGHVGVFVRAAIVCTRCVAKAVDFIPVES 1200
Qy 1201 METTMRSPVFTDNSTPPAVPQTFQVAHLHAPTGGSKSTKVPAAVAAQYKVLVNLNPSVAA 1260
Db 1201 METTMRSPVFTDNSTPPAVPQTFQVAHLHAPTGGSKSTKVPAAVAAQYKVLVNLNPSVAA 1260
Qy 1261 TLGFGAYMSKAHGDIPNRTIGVRTITTGGSITYSTYKGFLADGGCGSGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGDIPNRTIGVRTITTGGSITYSTYKGFLADGGCGSGAYDIIICDECHS 1320
Qy 1321 TDSITLIGITVLDQAEATAGARLVVLATATPPGSVTVPHPNIEBIGLSNNGEIEIPYGKAI 1380
Db 1321 TDSITLIGITVLDQAEATAGARLVVLATATPPGSVTVPHPNIEBIGLSNNGEIEIPYGKAI 1380
Qy 1381 PTEAIKGRHILIFCHSKKKCDLAAKLTGLGLNAVAYYRGLDVSVIPPIGDWWWATDAL 1440
Db 1381 PTEAIKGRHILIFCHSKKKCDLAAKLTGLGLNAVAYYRGLDVSVIPPIGDWWWATDAL 1440

Qy 1441 MTGFTGDFSDVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVRSORRGRTGRSGSIYR 1500
Db 1441 MTGFTGDFSDVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVRSORRGRTGRSGSIYR 1500
Qy 1501 FVTPGERPSGMFSSSVLCECYDAGCAWYELTPAETSURLRAYLNTPLGVPQODHLEFWES 1560
Db 1501 FVTPGERPSGMFSSSVLCECYDAGCAWYELTPAETSURLRAYLNTPLGVPQODHLEFWES 1560
Qy 1561 VFTGLTHIDAHFLSOTKQAGDNFPYLVAQATVCARAQAAPPSPWDQWKCLIRLKPILHG 1620
Db 1561 VFTGLTHIDAHFLSOTKQAGDNFPYLVAQATVCARAQAAPPSPWDQWKCLIRLKPILHG 1620
Qy 1621 PTPLLYRLGAVQNEVILTHPIIKYIMACMSADLEVVTSTWLVGVGLAALAAAYCLTTGVS 1680
Db 1621 PTPLLYRLGAVQNEVILTHPIIKYIMACMSADLEVVTSTWLVGVGLAALAAAYCLTTGVS 1680
Qy 1681 VIVGRIILSGKPAVVPDREVLYQEFDEMEECASQLPYIEQGMQLAEQFKOKALGILQAT 1740
Db 1681 VIVGRIILSGKPAVVPDREVLYQEFDEMEECASQLPYIEQGMQLAEQFKOKALGILQAT 1740
Qy 1741 KOEAAAAPVYESKWRALETFWAKHMNFISGIOYLAGLSTLPGNPAITASLMAFTASITSP 1800
Db 1741 KOEAAAAPVYESKWRALETFWAKHMNFISGIOYLAGLSTLPGNPAITASLMAFTASITSP 1800
Qy 1801 LTTQNTLLFNILGCVAAQALAPPASAFAVAGIAGAAVGSIGLGVKLVLDILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGCVAAQALAPPASAFAVAGIAGAAVGSIGLGVKLVLDILAGYGAGVA 1860
Qy 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEGAVOMNRLI 1920
Db 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEGAVOMNRLI 1920
Qy 1921 AFASRGHNHSPHYVPESDAAARVTQILSSLTITQLLKRLHOWINEBCSTPCSSGSLRDV 1980
Db 1921 AFASRGHNHSPHYVPESDAAARVTQILSSLTITQLLKRLHOWINEBCSTPCSSGSLRDV 1980
Qy 1981 WDMICTVLTDFKWLQSKLLPRLPGVPFLSCQYKGVWBGDGMOTTCPGAQIAGHVK 2040
Db 1981 WDMICTVLTDFKWLQSKLLPRLPGVPFLSCQYKGVWBGDGMOTTCPGAQIAGHVK 2040
Qy 2041 NGSMRIIVGPRTCNSTWHGTFFPINAYTTGCTPSPAPNYSRALMRVAAEYEVETRVGDFH 2100
Db 2041 NGSMRIIVGPRTCNSTWHGTFFPINAYTTGCTPSPAPNYSRALMRVAAEYEVETRVGDFH 2100
Qy 2101 YVTGMTTDNVKPCQVPAPPEFFTEVDGVRHLRVAPACKPLLRDVTTFQVGLNQYLVSQSL 2160
Db 2101 YVTGMTTDNVKPCQVPAPPEFFTEVDGVRHLRVAPACKPLLRDVTTFQVGLNQYLVSQSL 2160
Qy 2161 PCPEPEDVTVLTSMLTDPSSHITAETAKRRLARGSPPSLASSASQSLAPSILKATCTTHHD 2220
Db 2161 PCPEPEDVTVLTSMLTDPSSHITAETAKRRLARGSPPSLASSASQSLAPSILKATCTTHHD 2220
Qy 2221 SPDAADLLEANLLWRQEMGNITRVESENKVVILDSFEPLHAEGDEREISVAAEILRSRK 2280
Db 2221 SPDAADLLEANLLWRQEMGNITRVESENKVVILDSFEPLHAEGDEREISVAAEILRSRK 2280
Qy 2281 FPSALPIWARPDYNNPILLESWKDPDYVVPVHGCPLPTTKAPPTPPRRKRKTVVLTESNV 2340
Db 2281 FPSALPIWARPDYNNPILLESWKDPDYVVPVHGCPLPTTKAPPTPPRRKRKTVVLTESNV 2340
Qy 2341 SSALAEALATKTFGSSGSSAVDSGTATALPDLASDDGDKGDVESYSWMPLEGGPDPDL 2400
Db 2341 SSALAEALATKTFGSSGSSAVDSGTATALPDLASDDGDKGDVESYSWMPLEGGPDPDL 2400
Qy 2401 SDGSMSTVSEASDDVCCSMYSTWTGALITPCAABESKLPINPLNSLRLRHHNMVYATT 2460
Db 2401 SDGSMSTVSEASDDVCCSMYSTWTGALITPCAABESKLPINPLNSLRLRHHNMVYATT 2460
Qy 2461 SRASLRLQKKVTFDRLOVLDHVRDLVKEMKAKASTVAKLLSIEEACKLTPPHSAKSKF 2520
Db 2461 SRASLRLQKKVTFDRLOVLDHVRDLVKEMKAKASTVAKLLSIEEACKLTPPHSAKSKF 2520

Db 1 MGTNPKPQKTKNTNRRPODVKFGGQIVGGVYLLPRGPRGLGYRATRKTSERSQPRG 60
Qy 61 RRQPIPKARRPEGRAWAQPCGYPMWLYCNEGLGWAGMLLSPRGSRPSWGPTDPRRRSRNLG 120
Db 61 RRQPIPKARQPEGRAWAQPCGYPMWLYCNEGLGWAGMLLSPRGSRPSWGPTDPRRRSRNLG 120
Qy 121 KVIDTLTCGFADILMGVILPLVGA PLGGAARALAHGVRLVDGVNATCNLPSCFSIFLLA 180
Db 121 KVIDTLTCGFADILMGVILPLVGA PLGGAARALAHGVRLVDGVNATCNLPSCFSIFLLA 180
Qy 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIYEAADVIHMTFGVCPCVQBGNSRCWV 240
Db 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIYEAADVIHMTFGVCPCVQBGNSRCWV 240
Qy 241 ALPTTLAARNASVPTTIRRHVLLVGTAAFCAMVVDIGCGSIFLVSQIFTSPPRHET 300
Db 241 ALPTTLAARNASVPTTIRRHVLLVGTAAFCAMVVDIGCGSIFLVSQIFTSPPRHET 300
Qy 301 VQDNCNSIYPGHVSHRMAWDMNNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGLA 360
Db 301 VQDNCNSIYPGHVSHRMAWDMNNSPTTALVVSQLLRIPQAVVDMVAGAHGVLAGLA 360
Qy 361 YYSWVGWNAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKQLVNTNGSW 420
Db 361 YYSWVGWNAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKQLVNTNGSW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCCPERMASCRIPIWFAQGWGPITYTKPNSS 480
Db 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCCPERMASCRIPIWFAQGWGPITYTKPNSS 480
Qy 481 DQRPYCHYAPRCGVVPASQVCPVYCFTPSPVWVGTITDRSGVPTYSWGENETDVMULN 540
Db 481 DQRPYCHYAPRCGVVPASQVCPVYCFTPSPVWVGTITDRSGVPTYSWGENETDVMULN 540
Qy 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLICPTDCFRKHPEATYKCGSG 600
Db 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLICPTDCFRKHPEATYKCGSG 600
Qy 601 PWLTPRCLVDYPYRLMHPYCTLNFISI PKRMYVGVGEHRLNAAACNWTGERCNLEDRDS 660
Db 601 PWLTPRCLVDYPYRLMHPYCTLNFISI PKRMYVGVGEHRLNAAACNWTGERCNLEDRDS 660
Qy 661 ELSPLLSTTEWQILPCAFTTLPALSTGLIHLHONIVDVQVLYGVSAFVSFAIKWEYIL 720
Db 661 ELSPLLSTTEWQILPCAFTTLPALSTGLIHLHONIVDVQVLYGVSAFVSFAIKWEYIL 720
Qy 721 LIFLLADARVACCLMMWLLIAQAEALENLVINAASVAGAHGILSLFVFPFCAAWYIKG 780
Db 721 LIFLLADARVACCLMMWLLIAQAEALENLVINAASVAGAHGILSLFVFPFCAAWYIKG 780
Qy 781 RLAPGAAYFVGVNPLILLILLALPPRAYALDREMAASCAGVALVGLVFLTLSPYKVFILT 840
Db 781 RLVPGAAYLVGVNPLILLILLALPPRAYALDREMAASCAGVALVGLIILLTLSPHYKFLA 840
Qy 841 RLIMWLQVFIITRAEHAHQWVPPNLNVRGGRDAIILLTCVAPHELIIDITKULLAILGLPLM 900
Db 841 RLIMWLQVFIITRAEHAHQWVPPNLNVRGGRDAIILLTCVAPHELIIDITKULLAILGLPLM 900
Qy 901 VLQAGITVPYFVRAQGLIRACMLRVKAGHYVQWPMKLGALTGTVVYNHLLPLRDWA 960
Db 901 VLQAGITVPYFVRAQGLIRACMLRVKAGHYVQWPMKLGALTGTVVYNHLLPLRDWA 960
Qy 961 HAGRLDVAVEPVVFSAMETKVIITWAGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HAGRLDVAVEPVVFSAMETKVIITWAGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Qy 1021 QGWRLLAPITAYSQQTGRGLGCIITSLTGRDKNQVEGEVQVWSTATQSFATCINGVCWT 1080
Db 1021 QGWRLLAPITAYSQQTGRGLGCIITSLTGRDKNQVEGEVQVWSTATQSFATCINGVCWT 1080
Qy 1081 VYHGAGSKTLAGPKGPITOMYTNVDLVLGVHQAOPGARSMTFPCSCGSDLYLVTRHADVI 1140
Db 1081 VYHGAGSKTLAGPKGPITOMYTNVDLVLGVHQAOPGARSMTFPCSCGSDLYLVTRHADVI 1140

Qy 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGPILLCPSHVUVGFRAAACTRGVAKAVDFIPVES 1200
Db 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGPILLCPSHVUVGFRAAACTRGVAKAVDFIPVES 1200
Qy 1201 METTMESPVFTDNSTPPAVPQTFQVAHLHAPTCSGKSTKVPAAAYAAQGYKVLVLPNSVAA 1260
Db 1201 METTMESPVFTDNSTPPAVPQTFQVAHLHAPTCSGKSTKVPAAAYAAQGYKVLVLPNSVAA 1260
Qy 1261 TLGFGAYMSKAHGDINIRITGVNITITGSGIITVSTYTKGFLADGCGSGGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGDINIRITGVNITITGSGIITVSTYTKGFLADGCGSGGAYDIIICDECHS 1320
Qy 1321 TDSITTLIGTVDQOAEAGARLVVLATATPPGSVTVPHENIBIIGLSNNGEIPFYGKAI 1380
Db 1321 TDSITTLIGTVDQOAEAGARLVVLATATPPGSVTVPHENIBIIGLSNNGEIPFYGKAI 1380
Qy 1381 PIBAIIKGRHILIFCHSKKKCDELAALKITGLGLNAVAYYRGLDVSVIPPIGVVVVATDAL 1440
Db 1381 PIBAIIKGRHILIFCHSKKKCDELAALKITGLGLNAVAYYRGLDVSVIPPIGVVVVATDAL 1440
Qy 1441 MTGFTGDFSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVRSQRGRGTGRGRSIYR 1500
Db 1441 MTGFTGDFSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVRSQRGRGTGRGRMIYR 1500
Qy 1501 FVTPGERPSGMFSSVLCCEYDAGCAWYELTPAETSURLRAYLNTPGLPVCODHLEPWES 1560
Db 1501 FVTPGERPSGMFSSVLCCEYDAGCAWYELTPAETSURLRAYLNTPGLPVCODHLEPWES 1560
Qy 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAQYATVCARAQAAPPPSDQMWKCLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAQYATVCARAQAAPPPSDQMWKCLIRLKPTLHG 1620
Qy 1621 PTLLVRLGAVQNEVILTHPI TKYIMACMSADLEBVTSTWVLVGGVLAALAAAYCLITGTSV 1680
Db 1621 PTLLVRLGAVQNEVILTHPI TKYIMACMSADLEBVTSTWVLVGGVLAALAAAYCLITGTSV 1680
Qy 1681 VIVGRIILSKPAVVPDREVLVYOEFDMEERCASOLPYIEQGMQIAEOKKALGLQAT 1740
Db 1681 VIVGRIILSKPAIIPDREVLVYOEFDMEERCASOLPYIEQGMQIAEOKKALGLQAT 1740
Qy 1741 KQAEAAAAPVSVESKWRALETFWAKHMNFISGIOYLAGLSTLPGNPAPASLMAFTASITSP 1800
Db 1741 KQAEAAAAPVSVESKWRTEAFWAKHMNFISGIOYLAGLSTLPGNPAPASLMAFTASITSP 1800
Qy 1801 LTTQNTLLFNILGOWAAQALAPPASAFAVGAGIAGAAVSGIGLKVLDVILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGOWAAQALAPPASAFAVGAGIAGAAVSGIGLKVLDVILAGYGAGVA 1860
Qy 1861 GALVAFKVMGSEVPSTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGECAVOMNRLI 1920
Db 1861 GALVAFKVMGSEMPSTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGECAVOMNRLI 1920
Qy 1921 AFASRGHNVSTHVPSSDAAARVTQILSSLTITQLKRLHWINEDCSTPCSSGSLRDV 1980
Db 1921 AFASRGHNVSTHVPSSDAAARVTQILSSLTITQLKRLHWINEDCSTPCSSGSLRDV 1980
Qy 1981 WDMICTVLTDFKTLQSKLLPRLPGVFPFLSCQRYKGVWRGDGIMQITPCCGAQITGHVK 2040
Db 1981 WDMICTVLTDFKTLQSKLLPRLPGVFPFLSCQRYKGVWRGDGIMQITPCCGAQITGHVK 2040
Qy 2041 NGSMRIIVGPRCTSNWHTGTPPINAYTTGCTPPSPAPNYSRALWRVAAEYEVTRVGDPH 2100
Db 2041 NGSMRIIVGPRCTSNWHTGTPPINAYTTGCTPPSPAPNYSRALWRVAAEYEVTRVGDPH 2100
Qy 2101 YVTGWTNTDNVCKPCQVPAPEEFTVDGURLHRYAPACKPLLEDVTFVGLNQYLVSQOL 2160
Db 2101 YVTGWTNTDNVCKPCQVPAPEEFTVDGURLHRYAPACKPLLEDVTFVGLNQYLVSQOL 2160
Qy 2161 PCEPEPDVTVLTSMLTDPSSHITAETAKRRLARGSPPSLASSASQLSAPSKATCTTHHD 2220
Db 2161 PCEPEPDVTVLTSMLTDPSSHITAETAKRRLARGSPPSLASSASQLSAPSKATCTTHHD 2220

2221 SPADLIEANLWRQEMGNIIRVSENNKVWILDSFELHAEGRDREISVAAILRKRK 2280
Db 2221 SPADLIEANLWRQEMGNIIRVSENNKVWILDSFELHAEGRDREISVAAILRKRK 2280
Qy 2281 FPSALPIWARPDPYNPPLESWKDPDYPVPVHGCPDPPTKAPDIPPPRRKRTVUTESNV 2340
Db 2281 FPSALPIWARPDPYNPPLESWKDPDYPVPVHGCPDPPTKAPDIPPPRRKRTVUTESNV 2340
Qy 2341 SSALAEATKTPGSSGSAVDSTALPDLASDDGKSDVESYSSMPPLBEGPDPL 2400
Db 2341 SSALAEATKTPGSSGSAVDSTALPDQSDDDAGSDVESYSSMPPLBEGPDPL 2400
Qy 2401 SDGWSVTSEEAESDVCCSMSTWTGALITPCAAEESKLPINPLNSLLRHNNVYATT 2460
Db 2401 SDGWSVTSEEAESDVCCSMSTWTGALITPCAAEETKLPINALNSLLRHNNVYATT 2460
Qy 2461 SRSASLRQKKVTFDRQLQVLDHYRDVLEKMKAKASTVKAKLLSIEBAECKLTPPHSAKSF 2520
Db 2461 SRSASLRQKKVTFDRQLQVLDHYRDVLEKMKAKASTVKAKLLSIEBAECKLTPPHSAKSF 2520
Qy 2521 GYGAQDVRLNLSRAVNHHSVWEDLLEDTEPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAQDVRLNLSRAVNHHSVWEDLLEDTEPIDTTIMAKNEVFCVQPEKGRKPARLIV 2580
Qy 2581 FPDLGVRVCEKMAKYDVVSTLPOAVMGSSYQGYSPKQVFEFLVNTWKSCKCPMGFSYDT 2640
Db 2581 FPDLGVRVCEKMAKYDVVSTLPOAVMGSSYQGYSPQVFEFLVNAWKAKCPMGFAYDT 2640
Qy 2641 RCFDSTVTESDIRVESIYQCCDLAPEARQAIRSLTERLYIGPLTNSKGQNGYRRCKA 2700
Db 2641 RCFDSTVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGPLTNSKGQNGYRRCKA 2700
Qy 2701 SGVLTSCGNTLYCYLKATAACBAKLOCTMLVNGDDLVIWICESAGTOEDAAALRAFTE 2760
Db 2701 SGVLTSCGNTLYCYLKAAACBAKLOCTMLVCGDDLVIWICESAGTOEDASLRAFTE 2760
Qy 2761 AMTRYAPGDPPOPEYDLELITSCSNVSVAHADSGKRVYVLTDPPTPLARAAWETAR 2820
Db 2761 AMTRYAPGDPDPKPYDLELITSCSNVSVAHADSGKRVYVLTDPPTPLARAAWETAR 2820
Qy 2821 HTPINSWLGNIIMYAPTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQ 2880
Db 2821 HTPVNSWLGNIIMYAPTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQ 2880
Qy 2881 IERLHGLSAFTSHSYSGEINRVASCLRKLGVPPLRTWRHRARSVRKLLSQGGRAATC 2940
Db 2881 IIRLHGLSAFSLHSYSGEINRVASCLRKLGVPPLRVHRARSVRARLLSQGGRAATC 2940
Qy 2941 GRYLFNWAVRTKLTPIPAASOLDLSGWFWAGYSGDDIYHSLSRARPRWFPLCLLLSV 3000
Db 2941 GRYLFNWAVRTKLTPIPAASOLDLSWFWAGYSGDDIYHSLSRARPRWFPMWCLLLSV 3000
Qy 3001 GVGIVYLLPNR 3010
Db 3001 GVGIVYLLPNR 3010

RESULT 4
US-09-539-601-21
; Sequence 21, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlagler, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 3010

; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-21
Query Match 95.4%; Score 15279; DB 2; Length 3010;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2843; Conservative 80; Mismatches 87; Indels 0; Gaps 0;
Qy 1 MSTNPKPORKTKRNTNRRPQDVKYPGGQIVGGVYLLPRRGPRGLGVRATRKASERSQPG 60
Db 1 MGTNPKPORKTKRNTNRRPQDVKYPGGQIVGGVYLLPRRGPRGLGVRATRKATRSERSQPG 60
Qy 61 RRQPIPKARRPEGRAWAQPGYPWPLYGNEGIGWAGWLLSPRGSRPSWGPDPRRRSRLG 120
Db 61 RRQPIPKARQPEGRAWAQPGYPWPLYGNEGIGWAGWLLSPRGSRPSWGPDPRRRSRLG 120
Qy 121 KVIDTLTCGFADLMGYIPLVGCAPLGGARALAHGVRVLEEDGVNYATGNLPGCSFIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGCAPLGGARALAHGVRVLEEDGVNYATGNLPGCSFIFLLA 180
Qy 181 LLSCLTIPASAYEVNRSYGIYHVNTDCNSNIIYVEAADVIIMHTPGCVPCVOEGNSRRCW 240
Db 181 LLSCLTIPASAYEVNRSYGVTHVNTDCNSNIIYVEAADVIIMHTPGCVFCVRENNSSRCW 240
Qy 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYVGLDCSIFLVSQLFTFSRRHET 300
Db 241 ALTPTLAARNASVPTTIRRHVDLLVGAALCSAMVYVGLDCSIFLVSQLFTFSRRHET 300
Qy 301 VQDCNCSYIPGHVSHRMAMMMNWSPTTALVVSQLLRIPOAVVDMVAGAHGWGLAGLA 360
Db 301 VQDCNCSYIPGHVSHRMAMMMNWSPTTALVVSQLLRIPOAVVDMVAGAHGWGLAGLA 360
Qy 361 YYSWGNWAKVLIYVALLFAGVDGETHHTGRVAGHTTSGTSLFSSGASQKQLVNTNSGW 420
Db 361 YYSWGNWAKVLIYVALLFAGVDGGYVVGCTMAKNWLTGITSLSFSPSSQKQLVNTNSGW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCPERMASCRPIDWFAQMGPIYTKPNSS 480
Db 421 HINRTALNCNDSLOTGFLAALFYVHKFNSSGCPERMASCSPIDAFAQMGPIYTNESHSS 480
Qy 481 DORPYCMHYAPRPCGVVPASQVCGPVYCTPSVVPVVTDRSGVPTYSGENETDMLLN 540
Db 481 DORPYCMHYAPRPCGVVPAAQVCGPVYCTPSVVPVVTDRFPGVPTYSGENETDMLLN 540
Qy 541 NTRPPQGNWFCCTWNNSTGFTKTCGPPCNTGGVGNRTLIPTDCFRKHPEATYTKCSG 600
Db 541 NTRPPQGNWFCCTWNNSTGFTKTCGPPCNTGGVGNRTLIPTDCFRKHPEATYTKCSG 600
Qy 601 PMLTPRCCLVDYPYRLWHYPCTLNFISIFKVRMVYGVGVEHRLNAAACNWTGERCNLEDRS 660
Db 601 PMLTPRCCLVHYPYRLWHYPCTVNFIFKVRMVYGVGVEHRLNAAACNWTGERCNLEDRS 660
Qy 661 ELSPLLLSTTEWQILPCAFTTLPALSTGLIHLHQNIVDQVLYYGVGSAPVSAIKWEYIL 720
Db 661 ELSPLLLSTTEWQVLPSCFTTLPALSTGLIHLHQNIVDQVLYYGVGSAPVSAIKWEYIL 720
Qy 721 LFLILLADARVCACLWMLLIQAQAALLENLVNAAASVAGAHGILSLVFFCAAWYIKG 780
Db 721 LFLILLADARVCACLWMLLIQAQAALLENLVNAAASVAGAHGILSLVFFCAAWYIKG 780
Qy 781 RLAPGAAYAFYGVWPVLLLLLLALPPRAVALDREMAASCAGVAVLVFTLSPPYKVFIT 840
Db 781 RLVPGAAYALYGVWVFLLLLLALPPRAYAMDREMAASCAGVAVLVFTLSPPYKVFIT 840
Qy 841 RLIIWLOVPIITRAEAHMVWPVPLNVRGORDAILLTCVHPPELIPDITKLLAILGPLM 900
Db 841 RLIIWLOVPIITRAEAHLQWIPVPLNVRGORDAVILLTCAIHPELIFTITKLLAILGPLM 900
Qy 901 VLQAGITRVYPVFRAGGLIRACMLVRKAGGHYVQVFMKLCALTCYVYVNHLPDRWA 960
Db 901 VLQAGITKVPYFVRAGGLIRACMLVRKAGGHYVQVFMKLCALTCYVYVNHLPDRWA 960
Qy 961 HAGLRDLAVAVPVPVFSAMETKVTITWADTAACGDIIILGLPVSARRGKEIFLGPADSLEG 1020

Db 961 HAGLRDLAVAPVVFSDMETKVITWGAADTAACGDIIILGLPVSARRGREIHLGPADSLEG 1020
Qy 1021 QGWRLLAPIATAYSOOTRGVLGCIITSITGRDKNOVEGEVQVSTATQSFILATCINGVCWT 1080
Db 1021 QGWRLLAPIATAYSOOTRGVLGCIITSITGRDKNOVEGEVQVSTATQSFILATCINGVCWT 1080
Qy 1081 VYHGAGSKTLAGKGPITOMYTNVDLIVGQAPPGARSMTPCSCGSDILYLVTRHADVI 1140
Db 1081 VYHGAGSKTLAGKGPITOMYTNVDQDLVQWQAPPGARSLTPTCTCGSDILYLVTRHADVI 1140
Qy 1141 PVRRGDSRGLSLSPRVSYLKGSSGGPLLCPSHVVGVPRAAVCTRGVAKAVDFPVES 1200
Db 1141 PVRRGDSRGLSLSPRVSYLKGSSGGPLLCPSHVVGVPRAAVCTRGVAKAVDFPVES 1200
Qy 1201 METTMRSPVFTDNTSTPAVPTQFOVHLHAPTGGKSTKVPAAVAAQGYKVLVLPNSVAA 1260
Db 1201 METTMRSPVFTDNTSSPAVPTQFOVHLHAPTGGKSTKVPAAVAAQGYKVLVLPNSVAA 1260
Qy 1261 TLGFGAYMSKAHGDIPNIRITGGSTYTYGKFLADGGCGSGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGDIPNIRITGGSTYTYGKFLADGGCGSGAYDIIICDECHS 1320
Qy 1321 TDSTTILGIGVLDQAGTAGARLVVLTATPPGVTVPHPNIEIGLSNNGEIPFYGKAI 1380
Db 1321 TDSTTILGIGVLDQAGTAGARLVVLTATPPGVTVPHPNIEIGLSNNGEIPFYGKAI 1380
Qy 1381 PIEAKIGGRHLIFCHSKKKCDLAELTGLGLNAVAYYRGLDVSVIPIGDDVVVATDAL 1440
Db 1381 PIEAKIGGRHLIFCHSKKKCDLAELTGLGLNAVAYYRGLDVSVIPIGDDVVVATDAL 1440
Qy 1441 MTGTFGDFSDVIDCNTVQTVDFSLDPTFTIETTVPQDAVSQRGRGSGIYR 1500
Db 1441 MTGTFGDFSDVIDCNTVQTVDFSLDPTFTIETTVPQDAVSQRGRGSGIYR 1500
Qy 1501 FVTPGERPSGMDSSVLCEDYACAMVELTPAETSVELRAYLNTPLGVPQCDHLEFWS 1560
Db 1501 FVTPGERPSGMDSSVLCEDYACAMVELTPAETSVELRAYLNTPLGVPQCDHLEFWS 1560
Qy 1561 VFTGLTHIDAHFLSOTQOAGNFPYLVAYQATVCARAQAPPSPDQMWKCLIRLKTPLHG 1620
Db 1561 VFTGLTHIDAHFLSOTQOAGNFPYLVAYQATVCARAQAPPSPDQMWKCLIRLKTPLHG 1620
Qy 1621 PTPLLYRLGAYONEVILTHPTIKYIMACMSADLEVVTSTWLVGSLAALAAAYCLTTGSV 1680
Db 1621 PTPLLYRLGAYONEVILTHPTIKYIMACMSADLEVVTSTWLVGSLAALAAAYCLTTGSV 1680
Qy 1681 VIVGRILLSGKPAVVPDRVLYQEFDEMEECASOLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Db 1681 VIVGRILLSGKPAIIPDREVLYRPFDEMEECASHLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Qy 1741 KOAEAAAPVSVSKWRALETFWAKHMMNFIISGQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KOAEAAAPVSVSKWRTLEAFWAKHMMNFIISGQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Qy 1801 LTTONTLLFNILGWWAAQLAPPSAASAFVAGIAGAVSGIGLKVLDVILAGVGAVA 1860
Db 1801 LTTONTLLFNILGWWAAQLAPPSAASAFVAGIAGAVSGIGLKVLDVILAGVGAVA 1860
Qy 1861 GALVAFKVMSCVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQVMNRLI 1920
Db 1861 GALVAFKVMSCVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQVMNRLI 1920
Qy 1921 AFASRGNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQLHINEDCSTPCSGSLRDV 1980
Db 1921 AFASRGNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQLHINEDCSTPCSGSLRDV 1980
Qy 1981 WDWICTVLTFDKTWLQSKLLPRLPGVFLSCQRYKGVWRGDGIMQTTCPGQAQIAGHVK 2040
Db 1981 WDWICTVLTFDKTWLQSKLLPRLPGVFLSCQRYKGVWRGDGIMQTTCPGQAQIAGHVK 2040
Qy 2041 NGSMRIVGPRTCNTWHGTFFPINAYTTGCTPSPAPNYSRALWRVAEEYVEVTRVGDFH 2100

Db 2041 NGSMRIVGPRTCNTWHGTFFPINAYTTGCTPSPAPNYSRALWRVAEEYVEVTRVGDFH 2100
Qy 2101 YVTGMTNDNVKCPQVAPAPPEFFTEVDGVRHLRHRYAPACKPLLRBDVTPQVGLNQVLGSQL 2160
Db 2101 YVTGMTNDNVKCPQVAPAPPEFFTEVDGVRHLRHRYAPACKPLLRBEVTVLGLNQVLGSQL 2160
Qy 2161 PCPEPDVTVLTSMVLTDPHSHTAETAKRRLARGSPPSLASSASQSPSIPKATCTTHHD 2220
Db 2161 PCPEPDVTVLTSMVLTDPHSHTAETAKRRLARGSPPSLASSASQSPSIPKATCTTHHD 2220
Qy 2221 SPADLIIEANLLRWQMGNGNITRVESENKVVILDSFEPLHAEGDEREISVAAEILRSRK 2280
Db 2221 SPADLIIEANLLRWQMGNGNITRVESENKVVILDSFEPLHAEGDEREISVPAEILRSRK 2280
Qy 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
Db 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRERTVVLSBSTV 2340
Qy 2341 SSALAEIATKTFGSSGSAVDSTATATPDLASDDGDKGSDVESYSSMPPLLEGPGPD 2400
Db 2341 SSALAEIATKTFGSSGSAVDSTATATPDLASDDGDKGSDVESYSSMPPLLEGPGPD 2400
Qy 2401 SDGSWSVTSBEASEDVVCCSWSYTWGALITPCAAEBSKLPINPLNSLLRHHNMVYATT 2460
Db 2401 SDGSWSVTSBEASEDVVCCSWSYTWGALITPCAAEBSKLPINPLNSLLRHHNMVYATT 2460
Qy 2461 SRASLQKQKVTDFRLQVLDHVDVLEKEMKAKASTVKAKLLSTIEEACKLTPPHSAKSKF 2520
Db 2461 SRASLQKQKVTDFRLQVLDHVDVLEKEMKAKASTVKAKLLSTIEEACKLTPPHSAKSKF 2520
Qy 2521 GYGAKDVNRNLSRAVNHRSVWEDLLEDTETPIDTTIMAKSEVPCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVNRNLSRAVNHRSVWEDLLEDTETPIDTTIMAKSEVPCVQPEKGRKPARLIV 2580
Qy 2581 FPDGLGVCEKMAIYDVVSTLPQAVMGSSYGFQYSPKORVEFLVNTWKSKKCPMGESYDT 2640
Db 2581 FPDGLGVCEKMAIYDVVSTLPQAVMGSSYGFQYSPKORVEFLVNTWKSKKCPMGESYDT 2640
Qy 2641 RCFDSTVTESDIRVEESIYOCCLAPAEARQAIRSLTERLYIGGPLTNSKGONCGYRRCRA 2700
Db 2641 RCFDSTVTESDIRVEESIYOCCLAPAEARQAIRSLTERLYIGGPLTNSKGONCGYRRCRA 2700
Qy 2701 SGVLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICSSAGTQEDAAALRAPTE 2760
Db 2701 SGVLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICSSAGTQEDAAALRAPTE 2760
Qy 2761 AMTRYSAPPDPOPEVDLELITSCSSNVSAHDAKGRVYLLTRDPTTPPLARAWEATAR 2820
Db 2761 AMTRYSAPPDPOPEVDLELITSCSSNVSAHDAKGRVYLLTRDPTTPPLARAWEATAR 2820
Qy 2821 HTPINSWLGNTIMYAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQ 2880
Db 2821 HTPINSWLGNTIMYAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQ 2880
Qy 2881 IIERLHGLSAPLTHSYSGEINRVASCLRGVPLRTWRHRARSVRKLLSQGGRAATC 2940
Db 2881 IIERLHGLSAPLTHSYSGEINRVASCLRGVPLRTWRHRARSVRKLLSQGGRAATC 2940
Qy 2941 GRYLFNMAVTRKLTLPAPASQDLGSGWFGVAGYSGGDIYHSLGRARPRWFPCLLLLSV 3000
Db 2941 GRYLFNMAVTRKLTLPAPASQDLGSGWFGVAGYSGGDIYHSLGRARPRWFPCLLLLSV 3000
Qy 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010

RESULT 5
US-09-539-601-33
; Sequence 33, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW

! TITLE OF INVENTION: Hepatitis C Virus Cell Culture System

! FILE REFERENCE: all sequences

! CURRENT APPLICATION NUMBER: US/09/539,601C

! CURRENT FILING DATE: 2001-08-30

! EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY

! EARLIER FILING DATE: 1999-04-03

! NUMBER OF SEQ ID NOS: 51

! SOFTWARE: PatentIn Ver. 2.1

! SEQ ID NO 33

! LENGTH: 3010

! TYPE: PRT

! ORGANISM: Hepatitis C virus

! US-09-539-601-33

Query Match 95.4%; Score 15269; DB 2; Length 3010;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2841; Conservative 79; Mismatches 90; Indels 0; Gaps 0;

QY	1	MSNTNPKPQKTKRNTNRRPDQVKFPGGQIVGGVYLLPRRGPRLGVRATRKASERSQPRG	60
DB	1	MCTNPKPQKTKRNTNRRPDQVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG	60
QY	61	RRQPTPKARRPRGRAWAQQGYWPPLYGNEGLGWAGWLLSPRGRSPSWGPTDPRRNRNLG	120
DB	61	RRQPTPKARRPRGRAWAQQGYWPPLYGNEGLGWAGWLLSPRGRSPSWGPTDPRRNRNLG	120
QY	121	KVIDTLTCGFADLMGYIPLVGAAPLGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA	180
DB	121	KVIDTLTCGFADLMGYIPLVGAAPLGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA	180
QY	181	LLSCLTIPASAEYEVNRVSGIYHVNTDCNSNIIYEAADVIMETPGCPVCQEGNSRRCWV	240
DB	181	LLSCLTIPASAEYEVNRVSGIYHVNTDCNSNIIYEAADVIMETPGCPVCVRENNSRCWV	240
QY	241	ALTPTLAARNAVSPTTIRRHVDLLVGTAAFCSAMVVGDLGCSIFLVSOLFTEFSRRHET	300
DB	241	ALTPTLAARNAVSPTTIRRHVDLLVGAALCSAMVVGDLGCSVFLVAQLTEFSRRHET	300
QY	301	VQDCNCSIIYPGHVSGRHMAWMMNWSPTTALVVSOLLRIPOAVVDMVAGAHGVLGLA	360
DB	301	VQDCNCSIIYPGHVTGHRMAWMMNWSPTAALVVSOLLRIPOAVVDMVAGAHGVLGLA	360
QY	361	YYSWGNWAKVLIIVALLFAGVDEGTHTTGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW	420
DB	361	YYSWGNWAKVLIIVMLLIFAGVDEGTVTGTWAKNTLGTSLFSPSSQKIQLVNTNGSW	420
QY	421	HINRTALNCNDSIQTGFFAALFYAHKFNSSGCPERMASCRPIDWFAQGWGPITYKPNSS	480
DB	421	HINRTALNCNDSINTGFLAALFYVHKFNSSGCPERMASCSPIDAFAGQGWGPITYNESHS	480
QY	481	DORPYCWHYAPRCGVVPASOVCGPVYCTPSPVVVGTTRDGSVPYTSNGENETDVLIN	540
DB	481	DORPYCWHYAPRCGVVPAAQVCGPVYCTPSPVVVGTTRDFGVPTYSWGENETDVLIN	540
QY	541	NTRPPGNNFGCTWMNSTGFTKTCGGPPNCNIGVGNRTLICPTDCPRKHEPATYTKCGSG	600
DB	541	NTRPPGNNFGCTWMNSTGFTKTCGGPPNCNIGGINKTLTCTDCPRKHEPATYTKCGSG	600
QY	601	PMLTPRCLVDYPYRLWHYPCTLNFSIFKVRMYVGGVEHRLNAAACNWRGERCNLEDRDS	660
DB	601	PMLTPRCLVHPYRLWHYPCTVNFTIFKVRMYVGGVEHRLNAAACNWRGERCNLEDRDS	660
QY	661	ELSPILLSTTEWQILPCAFTTLPALSTGLIHONIVDVQYLVGVSASFVSAIKWEYIL	720
DB	661	ELSPILLSTTEWQVLPCEFTTLPALSTGLIHQNVVDVQYLVGVSASFVSAIKWEYVL	720
QY	721	LFLMLLADARVCACLWMMLLIAQEAALLENVLVNAASVAGAHGILSFLVFFCAAWYIKG	780
DB	721	LFLMLLADARVCACLWMMLLIAQEAALLENVLVNAASVAGAHGILSFLVFFCAAWYIKG	780
QY	781	RLAPGAAYAFYGVWPLLLLILLALPPRAYALDRMAASCGGAVLVGLVFTLTSPYKVLFT	840
DB	781	RLVPGAAYALYGVWPLLLLILLALPPRAYAMDRMAASCGGAVFVGLIILLTSPHYKFLA	840

QY	841	RLIWWLQYFITRAEAMQVWVPPNLNVRGRDAIILLTCAVHPELIFDITKLLAILGLPLM	900
DB	841	RLIWWLQYFITRAEAMQVWVPPNLNVRGRDAVILLTCAIHPELIFTITKILLAILGLPLM	900
QY	901	VLQAGITRVFVTRAQGLIRACMLVRKAGHYVOMFKMLGALTGTYYNHLTFLRDWA	960
DB	901	VLQAGITKVPYFVRAHGLIRACMLVRKAGHYVQMALMKLAALTGTYYDHLTFLRDWA	960
QY	961	HAGLRDLAVAVEPVVFSAMETKVIITWAGDATTACGDIILGLPVSARRGKEIFLGPADSLEG	1020
DB	961	HAGLRDLAVAVEPVVFSAMETKVIITWAGDATTACGDIILGLPVSARRGKEIHLGPADSLEG	1020
QY	1021	QWRLLEAPITAYSQOTRGVLGCIITSLTRCKDNQVEGEVQVYVSTATQSFATCINGVCWT	1080
DB	1021	QWRLLEAPITAYSQOTRGVLGCIITSLTRCKDNQVEGEVQVYVSTATQSFATCINGVCWT	1080
QY	1081	VYHGAGSKTLGAPGPIITOMYTNVDLDLVGWOAAPPGARSMTPCSCGSSDLYLVTRHADVI	1140
DB	1081	VYHGAGSKTLGAPGPIITOMYTNVDQDLVGWOAAPPGARSLTPTCTCGSSDLYLVTRHADVI	1140
QY	1141	PVRREGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAGVGFRAAVCTRGVAKAVDFVPVBS	1200
DB	1141	PVRREGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAGVGFRAAVCTRGVAKAVDFVPVBS	1200
QY	1201	METTNRSPVFTDNSTPPAVPQTFQVAHLHAPTSGSGKSTKVPAAYAAQGYKVLVLPNSVAA	1260
DB	1201	METTNRSPVFTDNSTPPAVPQTFQVAHLHAPTSGSGKSTKVPAAYAGQGYKVLVLPNSVAA	1260
QY	1261	TILGFGAYMSKAGHDIPNIRTVTTTGGSTIYTYSGKPLADGGCGGGAYDIIICDECHS	1320
DB	1261	TILGFGAYMSKAGHDIPNIRTVTTTGGSTIYTYSGKPLADGGCGGGAYDIIICDECHS	1320
QY	1321	TDSITLIGITVLDQAEATAGARLVLATATPGSVTVVPHNIEEIGLSNNGEIPFVGKAI	1380
DB	1321	TDSITLIGITVLDQAEATAGARLVLATATPGSVTVVPHNIEEIGLSNNGEIPFVGKAI	1380
QY	1381	PIEATIKGRHILFCHSKKKKDELAALTKGLGNLNAVAYRGLDVSVPITPDVWVATDAL	1440
DB	1381	PIETIKGRHILFCHSKKKKDELAALTKGLGNLNAVAYRGLDVSVPITPDVWVATDAL	1440
QY	1441	MTGFTGDFSDVSDCNCTVQTVDFSLDPTFTTETTTTPQDAVRSORRQRTGRGRGIYR	1500
DB	1441	MTGFTGDFSDVSDCNCTVQTVDFSLDPTFTTETTTTPQDAVRSORRQRTGRGRGIYR	1500
QY	1501	FVTPGERPSGMPDSSVLCEDYDAGCAYELTPTAETSRLRAYLNTPLGLPVCODHLEFWBS	1560
DB	1501	FVTPGERPSGMPDSSVLCEDYDAGCAYELTPTAETSRLRAYLNTPLGLPVCODHLEFWBS	1560
QY	1561	VFTGLTHIDAHFLSQTQKAGDNFPYLVAQATVCARAQAPPPSWDQMKCLIRLKPITLHG	1620
DB	1561	VFTGLTHIDAHFLSQTQKAGDNFPYLVAQATVCARAQAPPPSWDQMKCLIRLKPITLHG	1620
QY	1621	PTPLLYRLGAVQNEVILTHPIKYIMACMSADLEVVVSTWVLVGGVLAALAAAYCLTTGSV	1680
DB	1621	PTPLLYRLGAVQNEVITTHPIKYIMACMSADLEVVVSTWVLVGGVLAALAAAYCLTTGSV	1680
QY	1681	VIVGRIILISGRPAVDPDRVLYQBEDMEECASOLPYIEQGMQLAEQFKOKALGLLOTAT	1740
DB	1681	VIVGRIILISGRPAIIPDREVLRYEFDEMEECASHLPYIEQGMQLAEQFKOKAIGLLQATAT	1740
QY	1741	KQAEAAAAPVESKRALETFFWAKHWNFTSGIYQYLAGLSTLPGNPAIASIIMAFASITSP	1800
DB	1741	KQAEAAAAPVESKRWITIEAFWAKHWNFTSGIYQYLAGLSTLPGNPAIASIIMAFASITSP	1800
QY	1801	LTTQNTLLFNILGCVAAQALAPPASAAPVAGIAGAAVGSIGLGVLDIILAGYGAGVA	1860
DB	1801	LTTQNTLLFNILGCVAAQALAPPASAAPVAGIAGAAVGSIGLGVLDIILAGYGAGVA	1860
QY	1861	GALVAFKMSGEVPSSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQWNNRLI	1920
DB	1861	GALVAFKMSGBELPSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQWNNRLI	1920

QY 1921 AFASRGHVSPTHYVPESDAAARVTQILSSLTITQLLKRHLQWINEDCSTPCSGSWLRDV 1980
DB 1921 AFASRGHVSPTHYVPESDAAARVTQILSSLTITQLLKRHLQWINEDCSTPCSGSWLRDV 1980
QY 1981 WDWICTVLDTFKTWSKLLPLPGVPPFLSCQRYGKVGWRGDMOTTCPCGAQIAGHVK 2040
DB 1981 WDWICTVLDTFKTWSKLLPLPGVPPFLSCQRYGKVGWRGDMOTTCPCGAQIAGHVK 2040
QY 2041 NGSMRIVGPRCSTNTWHGTFPPINAYTTGCTPSPAPNYSALWRVAABEYVEVTRVGDHF 2100
DB 2041 NGSMRIVGPRCSTNTWHGTFPPINAYTTGCTPSPAPNYSALWRVAABEYVEVTRVGDHF 2100
QY 2101 YVTGMTDNDVKCQCPAPPEFFTEVDGRLHRYAPACKPLLRREDVTFQVGLNQLVGSOL 2160
DB 2101 YVTGMTDNDVKCQCPAPPEFFTEVDGRLHRYAPACKPLLRREDVTFQVGLNQLVGSOL 2160
QY 2161 PCPEPEPDVTLTSMULTDPSHITAEAKRLRLARGSPPLSSASSASQSLKATCTTHHD 2220
DB 2161 PCPEPEPDVTLTSMULTDPSHITAEAKRLRLARGSPPLSSASSASQSLKATCTTHHD 2220
QY 2221 SPADLIEANLLWRQMGNGNITRVESENKVILDSFEPLHAEGDEREISVAABILKRSK 2280
DB 2221 SPADLIEANLLWRQMGNGNITRVESENKVILDSFEPLHAEGDEREISVAABILKRSK 2280
QY 2281 FPSALPIWAPDYNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
DB 2281 FPSALPIWAPDYNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
QY 2341 SSALAEIATKTFSSGSSAVDSGTATLPDLASDDGDKGSDVESYSSMPPLEGEPPDIL 2400
DB 2341 SSALAEIATKTFSSGSSAVDSGTATLPDLASDDGDKGSDVESYSSMPPLEGEPPDIL 2400
QY 2401 SDGSWSVSEAEEDVCCSMYSYTWGALITPCAAEESKLPINPLSNLSLRLHNMVYATT 2460
DB 2401 SDGSWSVSEAEEDVCCSMYSYTWGALITPCAAEESKLPINPLSNLSLRLHNMVYATT 2460
QY 2461 SRASLRQKKVTFDLQVLDHRYDLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSKF 2520
DB 2461 SRASLRQKKVTFDLQVLDHRYDLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSKF 2520
QY 2521 GYGAKOVNLSRAVNHRSWEDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
DB 2521 GYGAKOVNLSRAVNHRSWEDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
QY 2581 FPDGLGVRCVCEKALYDVVSTLPOAVMGSSYGFQSPQKQVVEFLVNTWKSCKPCMGFSYDT 2640
DB 2581 FPDGLGVRCVCEKALYDVVSTLPOAVMGSSYGFQSPQKQVVEFLVNTWKSCKPCMGFSYDT 2640
QY 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNCYRRARA 2700
DB 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNCYRRARA 2700
QY 2701 SGVLTTCGNTLTCYLKATACRAAKLODCTMLVNGDDLVIICESAGTOEADAALRAFTE 2760
DB 2701 SGVLTTCGNTLTCYLKATACRAAKLODCTMLVNGDDLVIICESAGTOEADAALRAFTE 2760
QY 2761 AMTRYSAAPGPPPOPEYDLELITSCSNVSVVAHDASGRVYTLTRDPTTLARAAWETAR 2820
DB 2761 AMTRYSAAPGPPPOPEYDLELITSCSNVSVVAHDASGRVYTLTRDPTTLARAAWETAR 2820
QY 2821 HTPINSWLGNIIMVAPTLWARMILMTWFFSILLAQEQLKALDCQIYGACYSIEPLDLQP 2880
DB 2821 HTPINSWLGNIIMVAPTLWARMILMTWFFSILLAQEQLKALDCQIYGACYSIEPLDLQP 2880
QY 2881 IIERLHGLSATLHLSYSGEINRVASCLURKLGVPPLTRWRHARSVRKLLSQGGRAATC 2940
DB 2881 IIERLHGLSATLHLSYSGEINRVASCLURKLGVPPLTRWRHARSVRKLLSQGGRAATC 2940
QY 2941 GRYLFNNAVTRKLTPTIPASQDLSCWGFVAGYSGGDIYHSLSPARPRWFPCLILLUSV 3000
DB 2941 GRYLFNNAVTRKLTPTIPASQDLSCWGFVAGYSGGDIYHSLSPARPRWFPCLILLUSV 3000
QY 3001 GVGIVLLPNR 3010

DB 3001 GVGIVLLPNR 3010
RESULT 6
US-08-324-977-2
; Sequence 2, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-324-977-2

Query Match 95.3%; Score 15253; DB 1; Length 3010;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2840; Conservative 78; Mismatches 92; Indels 0; Gaps 0;

QY 1 MSTHPKQRTKRTNRRPQDKPFGGQIVGVYLLPRRGLGVRATRKASERSQPRG 60
|||||

Db 1 MSTNPKPQKTKRNTNRRPQDVKFGGQIVGVVYLLPRRGPRLGVRAPRKTSERSQPRG 60
Qy 61 RROPIPKARRPBGRAWAOPGYWPPLYGNEGLGWAQWLLSPRGRSPSWGPTDPRRRSNLG 120
Db 61 RKQPIPKARRPBGRTWAQPGYWPPLYGNEGLGWAQWLLSPRGRSPSWGPTDPRRRSNLG 120
Qy 121 KYIDTLTCGFADLMGYIPLVGAFLGAPGAARALAHGVVRVLEDGVYVATGNLPGCSFISFLLA 180
Db 121 KYIDTLTCGFADLMGYIPLVGAFLGAPGAARALAHGVVRVLEDGVYVATGNLPGCSFISFLLA 180
Qy 181 LLSCLTIPASAVEVRNVSIGYIYHVTDNCSNSIIVYBAADVIMHTPGCVPCVQEGNSRCWV 240
Db 181 LLSCLTTPASAYEVHNVSGIYHVTDNCSNASIIVYBAADLIHMTPGCVPCVQEGNSRCWV 240
Qy 241 ALTPTLAARNASVPPTTIRRHVDLLVGTAAFCSAMVYVGLDLCGSIFLVSQLTFSPRRHET 300
Db 241 ALTPTLAARNVTPITTTIRRHVDLLVGAAGFCSAMVYVGLDLCGSFVLVSQLTFSPRRHVT 300
Qy 301 VQDCNCSIYPGHVSGHRMAWDMMNWSPPTTALVWSQLLRIPOAVVDMVAGAHGWVLGLA 360
Db 301 LQDCNCSIYPGHVSGHRMAWDMMNWSPPTTALVWSQLLRIPOAVVDMVAGAHGWVLGLA 360
Qy 361 YYSVMGNNAKVLIVALLFAGVDGETHHTRGVAGHTTSGFTSPSSGASQIKIOLVNTNGSW 420
Db 361 YYSMAGNNAKVLIVMLLFAGVDGDTHVTGGAQKTTNRLVSMFASGPSQIKIOLVNTNGSW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSGCGPERMASCRPIDWFAQHGPIITYTKPNSS 480
Db 421 HINRTALNCNDSLOTGFLAALFYTHSFNSSGCGPERMAQOCRTIDKFDQMGPIITYAESRS 480
Qy 481 DORPYCWHYAPPCGVVPASQVCGPVYCFPTSPVGVVGTDRSGVPTYSGENETDVMLIN 540
Db 481 DORPYCWHYPPQCTIIVPASEVCGPVYCFPTSPVGVVGTDRSGVPTYRGENETDVMLIN 540
Qy 541 NTRPPQGNWFGCTWMNSTGFTKTCGPPCNIIGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Db 541 NTRPPQGNWFGCTWMNSTGFTKTCGPPCNIIGVGNNTLTCTDCFRKHPEATYTKCGSG 600
Qy 601 PMLTPRCLVDYPRLWHYPCITLNFSTIFKVRMYGVGVEHRLNAACNWRGERCNLEDRDRS 660
Db 601 PMLTPRCMVDPYPRLWHYPCVNTFTIFKVRMYGVGVEHRLNAACNWRGERCNLEDRDRP 660
Qy 661 ELSPLLSTTEWOILPCAFTTLPALSTGLIHLHQNIVDVQYLYGVGSAFVSFAIKWEYIL 720
Db 661 ELSPLLSTTEWOQLVPCSTFTPLPALSTGLIHLHQNIVDVQYLYGVGSAFVSFAIKWEYVL 720
Qy 721 LFLFLLLADARVCACLMWMLLIQAEEALENLVVLSASVAGAHGILSFLVFFCAAWYIKG 780
Db 721 LFLFLLLADARVCACLMWMLLIQAEEALENLVVLSASVAGAHGILSFLVFFCAAWYIKG 780
Qy 781 RLAPGNAAYAFYGVWPLLIILLLALPPRAYALDREMAASCAGAVLVGLVFTLSPYKYVELT 840
Db 781 RLVPGATYALYGVWPLLIILLLALPPRAYAMDREMAASCAGAVLVGLVFTLSPYKYVFLA 840
Qy 841 RLILWLOYFITRAEAHMOWVPPNVRGGRDAIILITCAVHPBELIFDITKLLAILGLPLM 900
Db 841 RLILWLOYFITTRAEDLHWIIPPLNARGGRDAIILMCAVHPBELIFDITKLLAILGLPLM 900
Qy 901 VLQAGITRPYFYFRAQGLIACMLVRKAGGHVQVMFKLGALETGYVYNHLTPLRDWA 960
Db 901 VLQAGITRPYFYFRAQGLIHACMLVRKAGGHVQVMAFMKLGALITGYIYNHLTPLRDWP 960
Qy 961 HAGRLDLAVABEVVPSAMETKVIWGAADTAACGDIILGLPVSARRGKEIFLGPADSLBG 1020
Db 961 RAGRLDLAVABEVVPSAMETKIIWGAADTAACGDIILGLPVSARRGKEIILGPADSLBG 1020
Qy 1021 QGWRLLAPITAYSQOQPRGLVGCIIITSLTRDKNOQVEGVQVWSTATQSFATCINGVCWT 1080
Db 1021 RGLRLLAPITAYSQOQTRGLVGCIIITSLTRDKNOQVEGVQVWSTATQSFATCINGVCWT 1080
Qy 1081 VYHGAGSKTLAGPKGITQMYTNVDLDLVGQWAPPGARSMTPCSGSSDLYLVTRHADVI 1140
Db 1081 VYHGAGSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSILTPTCGSSDLYLVTRHADVI 1140

Qy 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGLLPCPSGHVVGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGLLPCFPGHANGIFRAAVCTRGVAKAVDFIPVES 1200
Qy 1201 METTMRSPVFTDNSTPPPAVPQTFQVAHLHAPTGSKSTKVPAAYAAQYKVLVLPNSVAA 1260
Db 1201 METTMRSPVFTDNSTPPPAVPQSQFQVAHLHAPTGSKSTKVPAAYAAQYKVLVLPNSVAA 1260
Qy 1261 TLGFGAYMSKAHGIDPNIRTCVTRITTTGGSITYSTYSGFLADGGCGSGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGIDPNIRTCVTRITTTGAPTSTYSGFLADGGCGSGAYDIIICDECHS 1320
Qy 1321 TDSFTTILGIGTVLQDOAETAGARLVVLTATATPPGSVTVPHPNIEETGLNNGNEIPIFYGKAI 1380
Db 1321 TDSFTTILGIGTVLQDOAETAGARLVVLTATPPGSVTVPHPNIEEVALNNTGEIPIFYGKAI 1380
Qy 1381 PIEAIRKGRHILFCHSKKKKDELAALKTGLGLNAVAYYRGLDVSVIPPIGDVVVVATDAL 1440
Db 1381 PIEAIRKGRHILFCHSKKKKDELAALKSLGLINAVAYYRGLDVSVIPITIGDVVVVVATDAL 1440
Qy 1441 MTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVRSORRGRTGRGSGIYR 1500
Db 1441 MTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVRSORRGRTGRGSGIYR 1500
Qy 1501 FVTPGERPSGMFDSVLCCEYDAGCAWYELTPAETSVELRAYLNTPLPVCODHLEFPWES 1560
Db 1501 FVTPGERPSGMFDSVLCCEYDAGCAWYELTPAETSVELRAYLNTPLPVCODHLEFPWES 1560
Qy 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAQYATVCARAQPPSPDWQMKCLRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAQYATVCARAQPPSPDWQMKCLRLKPTLHG 1620
Qy 1621 PTPLLYRLGAVQNEVILTHPTIKYIMACNSADLEVVTTSTWLVGVVLAALAYCLTTGVS 1680
Db 1621 PTPLLYRLGAVQNEVILTHPTIKYIMACNSADLEVVTTSTWLVGVVLAALAYCLTTGVS 1680
Qy 1681 VIVGRIILSGPAPVVDREVLYQEFDEMEECASQLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Db 1681 VIVGRIILSGPAPVVDREVLYQEFDEMEECASQLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Qy 1741 KOEAAAAPVWESKWRALFTFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KOEAAAAPVWESKWRALFTFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Qy 1801 LTTQNTLLFNILGWVAAQLAPPSAASAFVAGIAGAAVSGSIGCKVLVDIILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGWVAAQLAPPSAASAFVAGIAGAAVSGSIGCKVLVDIILAGYGAGVA 1860
Qy 1861 GALVAFKVMGSEVPSTEDLVNLLPALISPGALVGVVCAAIILRRHVGPGEAGVQWNNRLI 1920
Db 1861 GALVAFKVMGSEVPSTEDLVNLLPALISPGALVGVVCAAIILRRHVGPGEAGVQWNNRLI 1920
Qy 1921 AFASRGNHVSPTHYVPESDAAAARVTQILSSLTITQLKRLHQMINEDECSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHYVPESDAAAARVTQILSSLTITQLKRLHQMINEDECSTPCSGSWLRDV 1980
Qy 1981 WDWICTVLITDFKTWLOSKLLPRLPGVPLSPCORGKYGVRGDMGIMQTTCCPQAGIAGHVK 2040
Db 1981 WDWICTVLITDFKTWLOSKLLPRLPGVPLSPCORGKYGVRGDMGIMQTTCCPQAGIAGHVK 2040
Qy 2041 NGSMEIVGPRTCNNTWHGTFPINAITYTTCPTSPAPNYSRALWRVAABEYVETVRGDFFH 2100
Db 2041 NGSMEIVGPRTCNNTWHGTFPINAITYTTCPTSPAPNYSRALWRVAABEYVETVRGDFFH 2100
Qy 2101 YVTGMTTNDNVKCPQVPAPEPFEVDGVRHLHRYAPACKPLLEDVTFQVGLNQYLVSQQL 2160
Db 2101 YVTGMTTNDNVKCPQVPAPEPFEVDGVRHLHRYAPACKPLLEDVTFQVGLNQYLVSQQL 2160
Qy 2161 PCEPEPDDVTLTSMULTDPSHITAETAKRRLARGSPPSLASSASQISAPSLKATCTTHHD 2220
Db 2161 PCEPEPDDVAVLTLTSMULTDPSHITAETAKRRLARGSPPSLASSASQISAPSLKATCTTHHV 2220

Qy	2221	SPADLIEANLLWRQEMGNITRVESENKVVILDSFEPLHAEGDEREISVAABILRKSRK	2288
Db	2221	SPADLIEANLLWRQEMGNITRVESENKVVILDSFDPLEAEEDREVSPAPBILRKSK	2288
Qy	2281	FPSALPIWARPDYNNPPLLESWKDPDYVPPVVHGCPLPPTKAPPIPPRRKRTVVLTESNV	2340
Db	2281	FPAAMPPIWARPDYNNPPLLESWKDPDYVPPVVHGCPLPPIKAPPIPPRRKRTVVLTESSV	2340
Qy	2341	SSALAEIATKTFGSGSSAVDSGTATALPLASDDGDKGSDVESYSMSMPLEGECPDPL	2400
Db	2341	SSALAEIATKTFGSESSAVDSGTATALPQASDDGDKGSDVESYSMSMPLEGECPDPL	2400
Qy	2401	SDGSWSTVSEASRDVVCCSWSYTWTCALITPCAABESKLPINPLNSLRLHHNMVYATT	2460
Db	2401	SDGSWSTVSEASEDDVVCCSWSYTWTCALITPCAABESKLPINALNSLRLHHNMVYATT	2460
Qy	2461	SRSASLRQKKVTFDRLOVLDDHYRDVILKEMKAKASTVKAKLLSIEBACKLTPPHSAKSKF	2520
Db	2461	SRSAGLRQKKVTFDRLOVLDDHYRDVILKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKF	2520
Qy	2521	GYGAKOVNRNLSRRAVNHIRSWEDLLEDTEPTDITTIMAKSEVFCVQPKGKPKPARLIV	2580
Db	2521	GYGAKOVNRNLSKAVNHIRHSWKDLLEDTEPTDITTIMAKNEVFCVQPKGKPKPARLIV	2580
Qy	2581	FPDLGVRVCEKMAIYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNTWKSXKCPMGFSYDT	2640
Db	2581	FPDLGVRVCEKMAIYDVVSTLPQVVMGSSYGFQYSPQORVEFLVNTWKSXKNPMGFSYDT	2640
Qy	2641	RCFDSVTSTSDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQONCGYRRCRA	2700
Db	2641	RCFPDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQONCGYRRCRA	2700
Qy	2701	SGVLTTSCGNLTLCYLKATACRAAKLQDCTMLVNGDDLVIVCESAGTQDEDAALRAFTE	2760
Db	2701	SGVLTTSCGNLTLCYLKASACRAAKLQDCTMLVNGDDLVIVCESAGTQDEDAASLRVFTE	2760
Qy	2761	AMTRYAPPDGPPOPEYDLELITSCSSNVSVAHDAAGKRVYVYLRDPTTTLAPAAWETAR	2820
Db	2761	AMTRYAPPDGPPOPEYDLELITSCSSNVSVAHDAAGKRVYVYLRDPTTTLAPAAWETAR	2820
Qy	2821	HTPINSWLGNTIMVAPTLWARMIWTHTFFSILLAQOLEKALDCQIYGACYSIEPLDLPQ	2880
Db	2821	HTPVNSWLGNTIMVAPTLWARMIWTHTFFSILLAQOLEKALDCQIYGACYSIEPLDLPQ	2880
Qy	2881	IIERLHGLSATHSYSGEINRVASCLRKLGVPPLRTWRHRSVRAKLLSQCGRAATC	2940
Db	2881	IIERLHGLSASFLSHSYSGEINRVASCLRKLGVPPLRVWRHRSVRAKLLSQCGRAATC	2940
Qy	2941	GXYLFNWAVRTKLKLTPIPAASOLDLSGWFFVAGYSGGDIYHLSLRARPFRWFLCLLLLSV	3000
Db	2941	GXYLFNWAVKTKLKLTPIPAASRLDLSGWFFVAGYSGGDIYHLSLRARPFRWFLCLLLLSV	3000
Qy	3001	GVGIYLLPNR 3010	
Db	3001	GVGIYLLPNR 3010	

RESULT 7
US-08-324-977-14
; Sequence 14, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton

[illegible]

Db	301	LQDCNCSIYPGHVSHRMAWMMNWSPTTALVVSQLLRIPQAVVDMVAGAHGWVLAGLA	360
Qy	361	YYSMVGNMAKVLIVALLFAGVDGETHHTTGRVAGHTTSFGTSLFSGASOKIOLVNTNGSW	420
Db	361	YYSMAGNMAKVLIVMLLFAGVDGDTHVTGGAQAKTINRLVSNFASGPSOKIOLINTNGSW	420
Qy	421	HINRTALNCNDSLOGTGFFAALFYAHKFNSSGCGPERMASCRPIDWFAQGWGPIITYTKPNSS	480
Db	421	HINRTALNCNDSLOGTGFLAALFYTHSFNSSGCGPERMAQCRTIDKPDQGWGPIITYAESSRS	480
Qy	481	DORPYCWHYAPRPGVVPASOVCGPVYCFTPBPVVVGTIDRSGVPTYSHGENETDVMLLN	540
Db	481	DORPYCWHYPPQCTIVPASEVCGPVYCFTPBPVVVGTIDRSGVPTYRGENETDVMLLN	540
Qy	541	NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIIGVGNRTLICPTDCFRKHPEATYTKCGSG	600
Db	541	NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIIGVGNNTLICPTDCFRKHPEATYTKCGSG	600
Qy	601	PWLTPRCLVDYPIYRLWHYPTCLNFSIFKVRMYVGVGVEHRLNAACNWTGERCNLEDRDRS	660
Db	601	PWLTPRCMVDPYRLWHYPTCVNFTIFKVRMYVGVGVEHRLNAACNWTGERCDEDRDRP	660
Qy	661	ELSPILLSTTEWQILPCAEFTTIPALSTGLIHLHONIVDVQYLYGYGSAFVSPAIKWEXYL	720
Db	661	ELSPILLSTTEWQVLPSCFTTIPALSTGLIHLHONIVDVQYLYGSAFVSPAIKWEXYL	720
Qy	721	LLFLLLLADARVCACLMWMLLIAQAEAALENLVVNAASVAGAHGILLSFLVFFCAAWYIKG	780
Db	721	LLFLLLLADARVCACLMWMLLIAQAEAALENLVVNSASVAGAHGILLSFLVFFCAAWYIKG	780
Qy	781	RLAPGAAYAFYGVWPILLLLALLPGRAYALDREMAASCGGAVLGVLFVTLSPYKVFELT	840
Db	781	RLVPGATYALYGVWPILLLLALLPGRAYAMDREMAASCGGAVFVGLVLLTSPYKVFELA	840
Qy	841	RLIWLQYFITRAEAMQWVPLNVRGORDAIILITCAVHPELIPDITKLLALLAILGPLM	900
Db	841	RLIWLQYFTTRAEADLHWIIPPLNARGGRDAIILIMCAVHPELIPDITKLLAILAILGPLM	900
Qy	901	VLQAGITRPYFVRAOGLIRACMLVRKVAGGHVQVFMKLGALTGTYVYNHLTPLROWA	960
Db	901	VLQAGITRPYFVRAOGLIHACMLVRKVAGGHVQVQAFMKLGALTGTYIYNHLTPLROWP	960
Qy	961	HAGRLDLAVAPVFPVFSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLG	1020
Db	961	RAGRLDLAVAPVFPVFSMETKIIITWGADTAACGDIILGLPVSARRGKEILLPADSLG	1020
Qy	1021	QWRLLAPITAYSQQTRGVLCGIITSLTGRDNQVGEVQVVSATQSFATCINGVCWT	1080
Db	1021	RGRLLLAPITAYSQQTRGLGCIITSLTGRDNQVGEVQVVSATQSFATCINGVCWT	1080
Qy	1081	VYHGAGSKTLAPKGPITQMYTNVDLDLVGWOAPPGARSMTPCSGSSDLYLVTRHADVI	1140
Db	1081	VYHGAGSKTLAPKGPITQMYTNVDQDLVGWPKPGARSMTPTCTCGSSDLYLVTRHADVI	1140
Qy	1141	PVRRGDSRGLISPRPVSYLKGSSGGLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES	1200
Db	1141	PVRRGDSRGLISPRPVSYLKGSSGGLLCPGHAVGIFRAAVCTRGVAKAVDFVPVES	1200
Qy	1201	METTWRSVFTDNSTPPAPQPTFOVAHLHAPTGSGHSTKVPAAYAAQGVKVLVLPNSVAA	1260
Db	1201	METTWRSVFTDNSSPPAVPQSFQVAHLHAPTGSGHSTKVPAAYAAQGVKVLVLPNSVAA	1260
Qy	1261	TLGFGAYMSKAGIDPNIETGVRTITTGSIINYSTYTKFLADGGCGGGAYDIIICDECHS	1320
Db	1261	TLGFGAYMSKAGIDPNIETGVRTITTGAPVTYSTYTKFLADGGCGGGAYDIIICDECHS	1320
Qy	1321	TDSTTILGTIVLDOAETAGARLVLATATPGSVTVPHPNTEIEIGLSNNGEIPFYGKAI	1380
Db	1321	TDSTTILGTIVLDOAETAGARLVLATATPGSVTVPHPNTEIEEVALSNTGEIPFYGKAI	1380
Qy	1381	PIEATKGGRHLLIFCHSKKKCDELAAKLTGLGLNANAYYRGLDVSVPPIPGDVVVVATDAL	1440
Db	1381	PIEATKGGRHLLIFCHSKKKCDELAAKLSGLGINANAYYRGLDVSVPPIPGDVVVVATDAL	1440
Qy	1441	MTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTVPQDAVRSQRRGRTGRSGIYR	1500
Db	1441	MTGYTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTVPQDAVRSQRRGRTGRSGIYR	1500
Qy	1501	FVTPGERPSGMFDSVLCCEVDAGCAWYELTTPAETSVRILRAYLNTPGLPVCOHDLHFWEBS	1560
Db	1501	FVTPGERPSGMFDSVLCCEYDAGCAWYELTTPAETSVRILRAYLNTPGLPVCOHDLHFWEBS	1560
Qy	1561	VFTGITHIDAFPLSOTKQAGNFPYLVAIQATVCARAQAPPSWDMKCLRLKPTLHG	1620
Db	1561	VFTGUTHIDAFPLSOTKQAGNFPYLVAIQATVCARAQAPPSWDMKCLRLKPTLHG	1620
Qy	1621	PTPLLRLGAVQNEVILTHPITKIMACMSADLEVVVTSTWLVGGVLAALAAAYCLTTGVS	1680
Db	1621	PTPLLRLGAVQNEVTLTHPITKIMACMSADLEVVVTSTWLVGGVLAALAAAYCLTTGVS	1680
Qy	1681	VIVGRILIISGKPAVVDREVLYQBEDMEBECASQPLYTEQGMQLAEOQKQKALGILLQAT	1740
Db	1681	VIVGRILIISGRPAIVPDRELLYQBEDMEBECASHLPYIEQGMQLAEOQKQKALGILLQAT	1740
Qy	1741	KOAEAAAAPVSESKWRALETFWAKHWNFIISGLOYLAGLSTLPGNPAIASLMAFTASITSP	1800
Db	1741	KOAEAAAAPVSESKWRALETFWAKHWNFIISGLOYLAGLSTLPGNPAIASLMAFTASITSP	1800
Qy	1801	LTTQNTLLFNILGGWVAQAAPPASAASAFVAGIAGAAVGSIGLGKVLVDIILAGYGAGVA	1860
Db	1801	LTTQSTLLFNILGGWVAQAAPPASAASAFVAGIAGAAVGSIGLGKVLVDIILAGYGAGVA	1860
Qy	1861	GALVAFKVMSEVSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEAGVQWNNRLI	1920
Db	1861	GALVAFKVMSEMPSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEAGVQWNNRLI	1920
Qy	1921	AFASGHNVSPTHYVPESDAAARVTOILSSLTITOLKRLHQMINEDECSTPCSGSWLRDV	1980
Db	1921	AFASGHNVSPTHYVPESDAAARVTOILSSLTITOLKRLHQMINEDECSTPCSGSWLRDV	1980
Qy	1981	WDWICTVLITDFKTLQSKLLPRLPGVPFLSCORGKGVWRGDGIMOTTCPGCAQTAGHVK	2040
Db	1981	WDWICTVLITDFKTLQSKLLPQLPGVPFLSCORGKGVWRGDGIMOTTCPGCAQITGHVK	2040
Qy	2041	NGSMRIVGPRTCSNTHGTFPFINAYTTGCTPSPAPNYSRALMRVAABEYEVTRVGDPH	2100
Db	2041	NGSMRIVGPRTCSNTHGTFPFINAYTTGCTPSPAPNYSRALMRVAABEYEVTRVGDPH	2100
Qy	2101	YVTGMTTNVCKPCQVPAPPEPTEVDGYLHRYAPACKPELLREDDVTFOGLNOYLVSQUL	2160
Db	2101	YVTGMTTNVCKPCQVPAPPEPTEVDGYLHRYAPACKPELLREDDVTFOGLNOYLVSQUL	2160
Qy	2161	PCEPEPDVTVLTSMULTDPSHITAETAKRRLARGSPPSLASSASCLSPSLKATCTTHHD	2220
Db	2161	PCEPEPDVAVLTSMULTDPSHITAETAKRRLARGSPPSLASSASQLSAPSLKATCTTHV	2220
Qy	2221	SPDADLIEANLWRQEMGNITRVESENKVILDSFEPHLHAEGBREISVAEILKRSK	2280
Db	2221	SPDADLIEANLWRQEMGNITRVESENKVVLDSFDPRLABEEDEREVSVAEILKRSK	2280
Qy	2281	PPSALPIHARDYNDPPILESWKDPDYVPVHVGCPLPTKAPPIPPPRKRTVVLTESNY	2340
Db	2281	PPAAMPPIHARDYNDPPILESWKDPDYVPVHVGCPLPIKAPPIPPPRKRTVVLTESV	2340
Qy	2341	SSALAEATKTFGSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPPLEGECPDPL	2400
Db	2341	SSALAEATKTFGSSSESAVDSGTATALPDQASDDGDKGSDVESYSSMPPLEGECPDPL	2400
Qy	2401	SDGWSVTSEASEDVCVCCSMSTWTGALIITPCAAEESKLPINPLSNSLLRHHNMVYAT	2460
Db	2401	SDGWSVTSEASEDVCVCCSMSTWTGALIITPCAAEESKLPINALSNSLLRHHNMVYAT	2460
Qy	2461	SRSASLROKVTTFDLQVLDHYRDLVKEMKAKASTVAKKLISIEEACKLTTPHSAKSKP	2520
Db	2461	SRSAGLROKVTTFDLQVLDHYRDLVKEMKAKASTVAKKLISVBEACKLTTPHSAKSKP	2520

2521 GYCAKVRNLSSRAVNHRSVWHELEDTETDTITIMAKSEVFCVQPEKGGKRPALIV 2580
2521 GYCAKVRNLSSRAVNHRSVWHELEDTETDTITIMAKSEVFCVQPEKGGKRPALIV 2580
2581 FPDGLVRVCEKALYDVVSTLPOAVMGSSYGFQSPKORVFLVNTWKSKKCPMGFSYDT 2640
2581 FPDGLVRVCEKALYDVVSTLPOAVMGSSYGFQSPKORVFLVNTWKSKKCPMGFSYDT 2640
2641 RCFDSTVTESDIRVESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGYRRCRA 2700
2641 RCFDSTVTESDIRVESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGYRRCRA 2700
2701 SGVLITSCGNTLTCYKATAACRAAKLOCTMLVNGDDLWVICSAGTOEDAAALRAFTE 2760
2701 SGVLITSCGNTLTCYKATAACRAAKLOCTMLVNGDDLWVICSAGTOEDAAALRAFTE 2760
2761 AMTRYSAPDGPDPPEYDLELITSCSSNVSVADHASKGVVYLTRDPTPLARAAMETAR 2820
2761 AMTRYSAPDGPDPPEYDLELITSCSSNVSVADHASKGVVYLTRDPTPLARAAMETAR 2820
2821 HTPINSWLGNIIIMYAPTLWARMILMTHFFSIIILAQBLEKALDCQIYGACYSIEPLDLPO 2880
2821 HTPINSWLGNIIIMYAPTLWARMILMTHFFSIIILAQBLEKALDCQIYGACYSIEPLDLPO 2880
2881 IIERLHGLSAFTLHVSPEINRVASCLIKGVPLRTWHRARSVRKLLSOGGGAATC 2940
2881 IIERLHGLSAFTLHVSPEINRVASCLIKGVPLRTWHRARSVRKLLSOGGGAATC 2940
2941 GRYLFNVAWTKLTPIPAASOLDLSGMFVAGYSGGDIYHLSLRARPRWFPLCLLLLSV 3000
2941 GRYLFNVAWTKLTPIPAASOLDLSGMFVAGYSGGDIYHLSLRARPRWFPLCLLLLSV 3000
3001 GVGIIYLLPNR 3010
3001 GVGIIYLLPNR 3010

RESULT 8
US-08-384-616-2
Sequence 2, Application US/08384616
Patent No. 5847101
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
ADDRESSEE: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-384-616-2
Query Match 95.3%; Score 15253; DB 1; Length 3010;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2840; Conservative 78; Mismatches 92; Indels 0; Gaps 0;
QY 1 MSTNPKPQKTKRNTNRRPQDVKPGGGQIVGGVILLPRGPRGLGVATRKASERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKPGGGQIVGGVILLPRGPRGLGVATRKASERSQPRG 60
QY 61 RRQPTPKARRPEGRAWAQPGYWPPLYGNEGLGWAGWLLSPRGSRPSWGTDPDRRRRNLG 120
Db 61 RRQPTPKARRPEGRAWAQPGYWPPLYGNEGLGWAGWLLSPRGSRPSWGTDPDRRRRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGARALAHGVRLVEDGVNVAATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGARALAHGVRLVEDGVNVAATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVYEAAADVIMHTPGVPCVQEGNSRCWV 240
Db 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVYEAAADVIMHTPGVPCVQEGNSRCWV 240
QY 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLCSGIFLVLSQIFTSRRHET 300
Db 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLCSGIFLVLSQIFTSRRHET 300
QY 301 VQDCNCSIYPGHVSGHRMADMMMNWSPITALVVSOLLRIPOAVDMVAGAHGVLGALA 360
Db 301 VQDCNCSIYPGHVSGHRMADMMMNWSPITALVVSOLLRIPOAVDMVAGAHGVLGALA 360
QY 361 YYSMVGWNAKVLIVALLFAGVDGETHHTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
Db 361 YYSMVGWNAKVLIVALLFAGVDGETHHTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
QY 421 HINRTALNCNDSLOTGFFAALFYAHKPNSSGCCPERMASCRPIDWFAQGCPITTKENSS 480
Db 421 HINRTALNCNDSLOTGFFAALFYAHKPNSSGCCPERMASCRPIDWFAQGCPITTKENSS 480
QY 481 DQRPYCHYHAPRCGVVPASOVGVYCFPTSPVAVVGTDRSGVPTYSWGENETDVMLLN 540
Db 481 DQRPYCHYHAPRCGVVPASOVGVYCFPTSPVAVVGTDRSGVPTYSWGENETDVMLLN 540
QY 541 NTRPPQGNWFCGCTWMNSTGFTKTCGGPPCNIGGVGNNTLTCPTDCFKKHEATYKCGSG 600
Db 541 NTRPPQGNWFCGCTWMNSTGFTKTCGGPPCNIGGVGNNTLTCPTDCFKKHEATYKCGSG 600
QY 601 PWLTPRCLVDYPYRLWHYPCTLNFSIPKVRMYGVGVEHRLNAACNWTGRRCNLEDKDRS 660
Db 601 PWLTPRCLVDYPYRLWHYPCTLNFSIPKVRMYGVGVEHRLNAACNWTGRRCNLEDKDRS 660
QY 661 ELSPLLLSTTEWQILLPCAFITLPAISGLIHLHQNIVDVQVLYGVGSFAFVSKWEVIL 720

[illegible][illegible]

QY 2881 IIERLHGLSFTLHVSYPGEINRVASCLRLKLGVPPLRTWHRARSVRAKLLSOGGRAATC 2940
Db 2881 IIERLHGLSFTLHVSYPGEINRVASCLRLKLGVPPLRVHRARSVRARLLSOGGRAATC 2940
QY 2941 GRYLFNVAWTKLKLTPIPAASOLDLSGWVAGVSGGDIYHLSLRARPRWFPLCLLLSV 3000
Db 2941 GKLYFNVAWTKLKLTPIPAASRLDLSGWVAGVSGGDIYHLSLRARPRWFPLCLLLSV 3000
QY 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010

RESULT 9

US-08-384-616-14
; Sequence 14, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FURE, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-384-616-14

Query Match 95.3%; Score 15253; DB 1; Length 3010;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2840; Conservative 78; Mismatches 92; Indels 0; Gaps 0;
QY 1 MSTNPKPQKRTKNTNRRPQDVKFGGQIVGGVYLLPRRGLRVRATRKASERSQPRG 60
Db 1 MSTNPKPQKRTKNTNRRPQDVKFGGQIVGGVYLLPRRGLRVRATRKASERSQPRG 60
QY 61 RROPIPKARRPEGRAWQFGYPMPLYGNEGLWAGWLLSPRGSPRSGWPTDPRRRSNLG 120
Db 61 RROPIPKARRPEGRTWAQFGYPMPLYGNEGLWAGWLLSPRGSPRSGWPTDPRRRSNLG 120
QY 121 KVIDTLTCGPADLGMGYIPLVGAPLGGAAALAHGVRLVDGVNATGNCPLGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLGMGYIPLVGAPLGGAAALAHGVRLVDGVNATGNCPLGCSFSIFLLA 180
QY 181 LLSCLTIPASAYEVNRVSGIYHVTNDCSNSSIYVEAADVIMHTPGVCVQOEGNSSRCWV 240
Db 181 LLSCLTIPASAYEVNRVSGIYHVTNDCSNASIYVEAADLIMHTPGVCVQOEGNSSRCWV 240
QY 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLGGISIFLVSQLFTFSPRHET 300
Db 241 ALPTTLAARNVTPTTIRRHVDLLVGAFAFCSAMVVDLGGISIFLVSQLFTFSPRHVT 300
QY 301 VQDCNCSIYPGHVSGHRMAMMMWNSPTTALVVSOLLRIPOAVVDMVAGAHGVLAGLA 360
Db 301 LQDCNCSIYPGHVSGHRMAMMMWNSPTTALVVSOLLRIPOAVVDMVAGAHGVLAGLA 360
QY 361 YYSWGVNWAQLIVALLFAGVDGETHTRGVAGHTSGFTSLFSSGASQIKLVNTNGSW 420
Db 361 YYSWAGNWAQLIVALLFAGVDGTHVTGGAQAKTNRLVSMFASGFSQIKLVNTNGSW 420
QY 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCGPERMASCRPIDWFAQGWGPITTKNSS 480
Db 421 HINRTALNCNDSLOTGFLAALFYTHSFSSGCGPERMAQCRITDKFDQGWGPITTYAESRS 480
QY 481 DQRPYCHYAPRCGVVPASQVCGPVYCFPTSPVVVCTTDRSGVPTYSWGENETDVMNLN 540
Db 481 DQRPYCHYPPQCTIVPASEVCGPVYCFPTSPVVVGTTRDFGVPTTRWGENETDVLN 540
QY 541 NTRPPQGNWFGCTWMNSTGFTKTCGPPCNIGGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Db 541 NTRPPQGNWFGCTWMNSTGFTKTCGPPCNIGGVGNRTLICPTDCFRKHPEATYTKCGSG 600
QY 601 PWTPLRCLVDYPYRLMHPYCTLNPSIFKVRMYVGVGVEHRLNAAACNWTGRCNLEDRDRS 660
Db 601 PWTPLRCMVDYPYRLMHPYCTVNFTIFKVRMYVGVGVEHRLNAAACNWTGRCNLEDRDRP 660
QY 661 ELSPLLLSTTEWQILPCAFITLPAALSTGLIHLHQNIVDVQVLYGVSFAFVSKWEYIL 720
Db 661 ELSPLLLSTTEWQVLPCTFTLPAALSTGLIHLHQNIVDVQVLYGVSFAFVSKWEYVL 720
QY 721 LFLLLADARVCACLMWMLLIAQAEALNVLNLAASVAGAHGILSLFVFFCAAWYIKG 780
Db 721 LFLLLADARVCACLMWMLLIAQAEALNVLNLAASVAGAHGILSLFVFFCAAWYIKG 780
QY 781 RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCGGAVLVGLVLTLSPPYKVELT 840
Db 781 RLVEGATYALYGVWPLLLLLALPPRAYAMDMAASCGGAVLVGLVLTLSPPYKVELA 840
QY 841 RLIIWLVQYFTRAEAHQVWVPPINVRGGRDAIILLCAVHPELIDITKLLALLGLPLM 900
Db 841 RLIIWLVQYFTRAEADLHVWIPPLNARGGRDAIILLCAVHPELIDITKLLIAILGLPLM 900
QY 901 VLQAGITRVYPYFVRAQGLIRACMLVRKVAGGHYVQVPMKLGALTGTVVVNHLPRLDWA 960
Db 901 VLQAGITRVYPYFVRAQGLIHACMLVRKVAGGHYVQVPMKLGALTGTVIYNHLPRLDWP 960
QY 961 HAGLRDLAVAVEPVVFSAMETKVTIWTGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 RAGLRDLAVAVEPVVFSAMETKIITWGNADTAACGDIILGLPVSARRGKEIILGPADSLEG 1020

1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDKNOQVEGEVQVSTATQSFATCINGVCWT 1080
1021 RGLRLAPITAYSQOTRGLGCIITSLTGRDKNOQVEGEVQVSTATQSFATCINGVCWT 1080
1081 VYHGAGSKTLAPKGPITOMYTNVNDLVLGVQAPFGARSMTPCSCGSSDLVLVTRHADVI 1140
1081 VYHGAGSKTLAPKGPITOMYTNVNDLVLGVQAPFGARSLTPTCTCGSSDLVLVTRHADVI 1140
1141 PVRRGDSRGSLLSRPVSYLKGSSGGLLCPSHVGVGFRAVCTRGVAKAVDFIPVES 1200
1141 PVRRGDSRGSLLSRPVSYLKGSSGGLLCPFHAVGIFRAVCTRGVAKAVDFIPVES 1200
1201 METTWRSVPFTDNTSPPAVPQTFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLPNSVAA 1260
1201 METTWRSVPFTDNTSPPAVPQSFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLPNSVAA 1260
1261 TLGFGAYMSKAHIGDIPNIRTGVRTITTTGGSITYSYGKPLADGGCGGAYDIIICDECHS 1320
1261 TLGFGAYMSKAHIGDIPNIRTGVRTITTTGAPVTYSTYGFADGGCGGAYDIIICDECHS 1320
1321 TDSITLIGTGLVLDQAETAGRLVLAATATPGSVTVPHPNIEEIGLSNNGEIPYVKAI 1380
1321 TDSITLIGTGLVLDQAETAGRLVLAATATPGSVTVPHPNIEEVALSNTGEIPYVKAI 1380
1381 PIEAIRGGRHLIFCHSKKCDLAAKLGLGLNAVAYRGLDVSIVPTIGDVVVVATDAL 1440
1381 PIEAIRGGRHLIFCHSKKCDLAAKLGLGINAVAYRGLDVSIVPTIGDVVVVATDAL 1440
1441 MTGFTGDFSDVDCNTCVTQTVDFSLDPTFTTETTVPODARSQRGRGTGRSGIYR 1500
1441 MTGFTGDFSDVDCNTCVTQTVDFSLDPTFTTETTVPODARSQRGRGTGRSGIYR 1500
1501 FVTPGERPMSGDFSVLCECYDAGCAWYELTAEYSVRLRAYINTPGLPVCOHLEFWES 1560
1501 FVTPGERPMSGDFSVLCECYDAGCAWYELTAEYSVRLRAYINTPGLPVCOHLEFWES 1560
1561 VFTGLTHIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSWDOMWKLRLKPTLHG 1620
1561 VFTGLTHIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSWDOMWKLRLKPTLHG 1620
1621 PTPLLYRLGAVQNEVLTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGVS 1680
1621 PTPLLYRLGAVQNEVLTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGVS 1680
1681 VIVGRILISGKPAVDREVLYQEFDEMEECASOLPYIEQGMQLAEOPKQKALGLLQAT 1740
1681 VIVGRILISGKPAVDREVLYQEFDEMEECASOLPYIEQGMQLAEOPKQKALGLLQAT 1740
1741 KQAEAAAPVVESKWRALFTFWAKHMWNFISGIQYLAGLSTLPGNPAIASIMAFITASITSP 1800
1741 KQAEAAAPVVESKWRALFTFWAKHMWNFISGIQYLAGLSTLPGNPAIASIMAFITASITSP 1800
1801 LTTQNTLLFNILGGWAAQLAPPSAASAFVAGIAGAAGVSGIGLKVLDVILLAGYGAGVA 1860
1801 LTTQNTLLFNILGGWAAQLAPPSAASAFVAGIAGAAGVSGIGLKVLDVILLAGYGAGVA 1860
1861 GALVAFKVMSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNRLLI 1920
1861 GALVAFKVMSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNRLLI 1920
1921 AFASRGNHVSPTHYVPESDAAARVTOILASLTTQLLKLHWINEDCSTPCSGSWLRDV 1980
1921 AFASRGNHVSPTHYVPESDAAARVTOILASLTTQLLKLHWINEDCSTPCSGSWLRDV 1980
1981 WDWICTVLTDFTKTLQSLPLRGLPVFPLSCQRYGKVGWRGDGIMOTTCPCGAQIAGHYK 2040
1981 WDWICTVLTDFTKTLQSLPLRGLPVFPLSCQRYGKVGWRGDGIMOTTCPCGAQIAGHYK 2040
2041 NGSMRIVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDHF 2100
2041 NGSMRIVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDHF 2100
2101 YVTGMTTDNVKPCQVPAPFEFFTEVDGVRLHRYAPACKPLLEDVTFQVGLNQYLVGSOL 2160

2101 YVTGMTTDNVKPCQVPAPFEFFSEVDGVRLHRYAPACKPLLEDVTFQVGLNQYLVGSOL 2160
2161 PCPEPDDVTVLTMLTDPSSHITAETAKRRLARGSPPSLASSASASOLGAPSKATCTTHHD 2220
2161 PCPEPDDVAVLUTSMLTDPSSHITAETAKRRLARGSPPSLASSASASOLGAPSKATCTTHHV 2220
2221 SPDADLIEANLLWRQENGMNITRVESENKVVILDSFEPFLAEGDEREISVAAEILRKSJK 2280
2221 SPDADLIEANLLWRQENGMNITRVESENKVVILDSFDPFLAEEDEEREVSVAEILRKSJK 2280
2281 PPSALPIWARPDYNNPPLLESWKDDYVPPVWHGCPCLPPTKAPPIPPPRKRKTIVLTESNV 2340
2281 PPAAMPPIWARPDYNNPPLLESWKDDYVPPVWHGCPCLPPIKAPPIPPPRKRKTIVLTESNV 2340
2341 SSALAEATKTFGSGSSAVDSGTATAPDLASDDGDKGSDVESYSSMPPLEGPDPDL 2400
2341 SSALAEATKTFGSGSSAVDSGTATAPDQASDDGDKGSDVESYSSMPPLEGPDPDL 2400
2401 SDGWSVTSEASDVVCCSMSTWTGALITPCCAABESKLPINPLNSLLRHHNMVYATT 2460
2401 SDGWSVTSEASDVVCCSMSTWTGALITPCCAABESKLPINALNSLLRHHNMVYATT 2460
2461 SRSASLRQKQVTFDRLQVLDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSF 2520
2461 SRSAGLRQKQVTFDRLQVLDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSF 2520
2521 GYGAKDVNLSRRVNHRSVWEDLLEDTEPTIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
2521 GYGAKDVNLSRRVNHRSVWEDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIV 2580
2581 FPDGLVRVCEKMAKYDVVSTLTPQAVMGSSYGFQYSPKQVBFVNTWTKSKCPCMGFSYDT 2640
2581 FPDGLVRVCEKMAKYDVVSTLTPQAVMGSSYGFQYSPQVBFVNTWTKSKKPMGFSYDT 2640
2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIKSLTERLYTGGPLTNSKGQNCVYRRCRA 2700
2641 RCFDSTVTENDIRVEESIYQCCDLAPEARQAIKSLTERLYTGGPLTNSKGQNCVYRRCRA 2700
2701 SGVLTSTCGNTLTCLYKATAACRAKLODCTMLVNGDDLVVICESAGTQEDAAALRAFTE 2760
2701 SGVLTSTCGNTLTCLYKASAAACRAKLODCTMLVNGDDLVVICESAGTQEDAAASURVFE 2760
2761 AMTRYASAPPPPOPEYDLELITSCSSNVSAHADSKRKYVYLTDRPTTPLARAAWETAR 2820
2761 AMTRYASAPPPPOPEYDLELITSCSSNVSAHADSKRKYVYLTDRPTTPLARAAWETAR 2820
2821 HTPINSWLGNTIMVAPTILWARMILMTHPFSILLAQOEKALDCQIYGACYSIEPLDLPQ 2880
2821 HTPVNSWLGNTIMVAPTILWARMILMTHPFSILLAQOEKALDCQIYGACYSIEPLDLPQ 2880
2881 IIERLHGLSAPTLHSYSGEINRVASCLKGLGVPLRTWRHRARSVRAKLLSQGGRAATC 2940
2881 IIERLHGLSAPTLHSYSGEINRVASCLKGLGVPLRTWRHRARSVRAKLLSQGGRAATC 2940
2941 GRYLFNVAVRTKLTLPPIAASQDLSCGFVAGYSGGDIYHSLSRARPRWFPLCLLLLSV 3000
2941 GRYLFNVAVRTKLTLPPIAASQDLSCGFVAGYSGGDIYHSLSRARPRWFPLCLLLLSV 3000
3001 GVGIYLLPNR 3010
3001 GVGIYLLPNR 3010

RESULT 10

US-08-904-686A-2

; Sequence 2, Application US/08904686A

; Patent No. 5981130

; GENERAL INFORMATION:

; APPLICANT: OKAYAMA, Hiroto

; APPLICANT: FUKU, Isao

; APPLICANT: MORI, Chisato

; APPLICANT: TAKAMIZAWA, Akahisa

APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westernman, Hattori, McLeeland &
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/904,686A

FILING DATE: 01-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/324,977

FILING DATE: 18-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-167466

FILING DATE: 25-JUN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-230921

FILING DATE: 31-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-305605

FILING DATE: 09-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/099,706

FILING DATE: 30-JUL-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/769,996

FILING DATE: 02-OCT-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/635,451

FILING DATE: 28-DEC-1990

ATTORNEY/AGENT INFORMATION:

NAME: McLeeland, Le-Nhung

REGISTRATION NUMBER: 31,541

REFERENCE/DOCKET NUMBER: 900703G

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 659-2930

TELEFAX: (202) 887-0357

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 3010 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-904-686A-2

Query Match 95.3%; Score 15253; DB 1; Length 3010;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2840; Conservative 78; Mismatches 92; Indels 0; Gaps 0;

Qy 1 MSTNPKQKTRNTRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATRKASERSQPRG 60
Db 1 MSTNPKQKTRNTRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATRKASERSQPRG 60

Qy 61 RRQIPKARRPEGRWAQPGYFWPLYGNEGLWAGWLLSPRGRSPSWGFTDPRRSRLG 120
Db 61 RRQIPKARRPEGRWAQPGYFWPLYGNEGLWAGWLLSPRGRSPSWGFTDPRRSRLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAAPLGAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAAPLGAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTTPASAYEVNRVNSGIYHVTNDCNSNIIYVEAADVIMHTPGCVPCVQEGNSSRCWV 240

Db 181 LLSCLTTPASAYEVNRVNSGIYHVTNDCNSNIIYVEAADVIMHTPGCVPCVQEGNSSRCWV 240
Qy 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAAFCSAMTVGDLCGSIFLVSQLFTSPRRHET 300
Db 241 ALTPTLAARNVPTTIRRHVDLLVGAFAAFCSAMTVGDLCGSVFLVSQLFTSPRRHVT 300
Qy 301 VQDCNCSIYPGHVSGHRMADMMNWSPTTALVVSOLLRIPOAVVDMVAGAHGVLAGLA 360
Db 301 LQDCNCSIYPGHVSGHRMADMMNWSPTTALVVSOLLRIPOAVVDMVAGAHGVLAGLA 360
Qy 361 YYSMVGNWAKVLIALLFAGVDGTHTCGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW 420
Db 361 YYSMAGNWKVLIIVMLLFAGVDGTHVTGGAQAKTNRVLSVMEFASGPSOKIQLINTNGSW 420
Qy 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCCPERMASCRPIDWPAQGWGPIITYRENSS 480
Db 421 HINRTALNCNDSLOTGFLAALFYTHSFSSGCCPERMAQCRITIDKFDQGWGPIITYAESSRS 480
Qy 481 DQRPYCHWYAPRPGCVVPASOVCGPVYCTPTSPVVVGTTRDSGVPTYSGENETDVMLLN 540
Db 481 DQRPYCHWYPPQCTIIPASEVCGPVYCTPTSPVVVGTTRDFRGVPTYRWGENETDVLNLLN 540
Qy 541 NTRPPQGNMFGCTWMNSTGFTKCGPPPCNIGVGNRTLCPTDCFRKHPEATVTKCGSG 600
Db 541 NTRPPQGNMFGCTWMNSTGFTKCGPPPCNIGVGNRTLCPTDCFRKHPEATVTKCGSG 600
Qy 601 PMLTPRCLVDYPYRLMHPCTLNFSIFKVRMYGVGVEHRLNAACNWTGRGERCNLEDRDRS 660
Db 601 PMLTPRCWVDYPYRLMHPCTVNFITFKVRMYGVGVEHRLNAACNWTGRGERCOLEDRDRP 660
Qy 720 ELSPELLSTTEWQILPCAFITLLPALSTGLIHLHQNIVDVQVLYGVGSFAFKWEVYL 720
Db 720 ELSPELLSTTEWQVLPCTFTLLPALSTGLIHLHQNIVDVQVLYGVGSFAFKWEVYL 720
Qy 780 LLFLLADARVCACLMWMLLIAQAEALLENVLVNAASVAGAHGILSFLVFFCAAWIYK 780
Db 780 LLFLLADARVCACLMWMLLIAQAEALLENVLVNSASVAGAHGILSFLVFFCAAWIYK 780
Qy 840 RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCAGAVLVGLVLTLSPPYKPLT 840
Db 840 RLVPGATYALYGVWPLLLLLALPPRAYANDREMAASCAGAVLVGLVLTLSPPYKFLA 840
Qy 900 RLIWMLQYFTRAEAHMNVVPLNVRGGRDAIILLTCAVHPELIIDITKLLAILGLPLM 900
Db 841 RLIWMLQYFTRAEADLHVNPPLNARGGRDAIILLMCAVHPELIIDITKLLAILGLPLM 900
Qy 960 VLQAGITRVYFVRAQGLIHACMLVRKVAGHYVQVFMKLGALTGTVYVNHLPRLDWA 960
Db 901 VLQAGITRVYFVRAQGLIHACMLVRKVAGHYVQVFMKLGALTGTIYVNHLPRLDWP 960
Qy 1020 HAGLRDLAVAVEPVVFSAMETKVITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 RAGLRDLAVAVEPVVFSAMETKVITWGADTAACGDIILGLPVSARRGKEILLGPADSLEG 1020
Qy 1080 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDNQVEGEVQVVSSTATQSLATCINGVCWT 1080
Db 1021 RGLRLLAPITAYSQOTRGVLGCIITSLTGRDNQVEGEVQVVSSTATQSLATCINGVCWT 1080
Qy 1081 VYHAGAGSKTLAPKGPITOMYTNVDLIVGHQAAPGARSMTPCSGSSDLVLTTRHADVI 1140
Db 1081 VYHAGAGSKTLAPKGPITOMYTNVDLIVGHQAAPGARSMTPCSGSSDLVLTTRHADVI 1140
Qy 1141 PVRRGRSRLSPRVSYLKGSSGGLLCPSGHVVGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRGRSRLSPRVSYLKGSSGGLLCPFGHAGVIFRAAVCTRGVAKAVDFIPVES 1200
Qy 1260 METTMRSPVFDNSTPPAVPQTFQVAHLHAPTGSKSTKPAAYAAQYKVLVLPNSVAA 1260
Db 1201 METTMRSPVFDNSTPPAVPQSFQVAHLHAPTGSKSTKPAAYAAQYKVLVLPNSVAA 1260
Qy 1261 TLGFGAYMSKAHGDIPNIRTGRTITTCGSIITYSTYTGKFLADGGCGSGAYDIILICDECHS 1320

Db 1261 TLGFGAYMSKAHIDPNIRITGVRTITTCAPVTYSTYKFLADGGCGGAYDIIICDECHS 1320
Qy 1321 TDSTIILGIVLDOAETAGARLVVLATATPPGSVTVPHPNIEEIGLNGNEIIPPYGKAI 1380
Db 1321 TDSTIILGIVLDOAETAGARLVVLATATPPGSVTVPHPNIEEVALSNTGEIIPPYGKAI 1380
Qy 1381 PTEALIKGRHLIFCHSKKKCDLAAKLTCGLGNVAVYVYRGLDVSVIPIGDVVVVATDAL 1440
Db 1381 PTEALIKGRHLIFCHSKKKCDLAAKLTCGLGNVAVYVYRGLDVSVIPIGDVVVVATDAL 1440
Qy 1441 MTGFTGDFSDVDCNTCVTQTQVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGSIYR 1500
Db 1441 MTGFTGDFSDVDCNTCVTQTQVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGSIYR 1500
Qy 1501 FVTPGERSGMPDSSVLCEDYDAGCAWYELTPAETSRLRAYLINTPGLPVCDHLEFWES 1560
Db 1501 FVTPGERSGMPDSSVLCEDYDAGCAWYELTPAETSRLRAYLINTPGLPVCDHLEFWES 1560
Qy 1561 VFTGLTHIDAHFLSOTKQAGDNPPYLVAVOATVCARAQAPPSWDMKCLIRLKPRTLHG 1620
Db 1561 VFTGLTHIDAHFLSOTKQAGDNPPYLVAVOATVCARAQAPPSWDMKCLIRLKPRTLHG 1620
Qy 1621 PTPLLYRLGAVQNEVILTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLITGVS 1680
Db 1621 PTPLLYRLGAVQNEVILTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLITGVS 1680
Qy 1681 VIVGRILLSGKPAVVPDREVLVQEFDEMEECASQLPYIEQGHQLABQFKQKALGILLQAT 1740
Db 1681 VIVGRILLSGKPAVVPDREVLVQEFDEMEECASQLPYIEQGHQLABQFKQKALGILLQAT 1740
Qy 1741 KQAEAAAPVVESKRALETFWAKHWNFIISGQYLAGLSTLPGNPAIASLMFTAFTSITSP 1800
Db 1741 KQAEAAAPVVESKRALETFWAKHWNFIISGQYLAGLSTLPGNPAIASLMFTAFTSITSP 1800
Qy 1801 LTTQNTLLFNILGGWVAQAAPPASAFAVGGIAGAAVGSICLGKVLVDIILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGGWVAQAAPPASAFAVGGIAGAAVGSICLGKVLVDIILAGYGAGVA 1860
Qy 1861 GALVAPKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNRLLI 1920
Db 1861 GALVAPKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNRLLI 1920
Qy 1921 AFASRGNHVSPTHYVPESDAARVTOILSSLTITOLLKELHOWINEDCSTPCSGWLRDV 1980
Db 1921 AFASRGNHVSPTHYVPESDAARVTOILSSLTITOLLKELHOWINEDCSTPCSGWLRDV 1980
Qy 1981 WDWICTVLDTFTWLSKLLPRLPVFPLSCORGKGVWRGDIQMTTCPCGAQIAGHVK 2040
Db 1981 WDWICTVLDTFTWLSKLLPQLPGVPFFCQGYKGVWRGDIQMTTCPCGAQITGHVK 2040
Qy 2041 NGSMRIVGPKTCSNTWHGTFPINAFTGCTPSPAPNYSRALWRVAAEYVEVTRVGDPH 2100
Db 2041 NGSMRIVGPKTCSNTWHGTFPINAFTGCTPSPAPNYSRALWRVAAEYVEVTRVGDPH 2100
Qy 2101 VYTGMTDNVCKPCQVPAPEFFTEVDGVRHLRVAPACKPLREDVTFOVLNQYLVSOL 2160
Db 2101 VYTGMTDNVCKPCQVPAPEFFTEVDGVRHLRVAPACKPLREDVTFOVLNQYLVSOL 2160
Qy 2161 PCEPEPDVLTLSMLTDPHSHTAETAKRLARGSPSLASSASQSLAPSLKATCTTHHD 2220
Db 2161 PCEPEPDVLTLSMLTDPHSHTAETAKRLARGSPSLASSASQSLAPSLKATCTTHV 2220
Qy 2221 SPDADLIEANLWRQEMGNITRVESENKVILDSPEPLHAEGDEREISVAEILRKSRK 2280
Db 2221 SPDADLIEANLWRQEMGNITRVESENKVILDSPEPLHAEGDEREISVAEILRKSRK 2280
Qy 2281 FFSALPIWARPDPNPNLLESKDPDVPVPHVGCPLPPTKAPPIPPRKRKTVLVTESNV 2340
Db 2281 FPAAMPWARPDPNPNLLESKDPDVPVPHVGCPLPPTKAPPIPPRKRKTVLVTESV 2340
Qy 2341 SSALAEATKTFGSSGSSAVDSGTATLPDLASDDGDKGSDVSYSSMPPLGEGDPDL 2400
Db 2341 SSALAEATKTFGSSSVAVDSGTATLPDLASDDGDKGSDVSYSSMPPLGEGDPDL 2400

Qy 2401 SDGSWSTVSEASDVVCCSMSTWTGALITPCAABESKLPINPLSNLRLHNNMVYATT 2460
Db 2401 SDGSWSTVSEASDVVCCSMSTWTGALITPCAABESKLPINALSNLRLHNNMVYATT 2460
Qy 2461 SRSASLRQKVTDFRLQVLDHDDYDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSF 2520
Db 2461 SRSASLRQKVTDFRLQVLDHDDYDLVKEMKAKASTVKAKLLSVEEACKLTPPHSAKSF 2520
Qy 2521 GYGAKDVRLNLSRAVNHRSWEDLLEDTEPIDTITIMAKSEVFCVQEKGRKPARLIV 2580
Db 2521 GYGAKDVRLNLSKAVNHRSWVKOLLEDTEPIDTITIMAKNEVFCVQEKGRKPARLIV 2580
Qy 2581 FPDGLGVRCEKALYDVVSTLPOAVMGSSYGFQYSPKORVBEFLVNTWKSCKCPMGFSYDT 2640
Db 2581 FPDGLGVRCEKALYDVVSTLPOAVMGSSYGFQYSPQORVBEFLVNTWKSCKNPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQCGYRRCRA 2700
Db 2641 RCFDSTVTENDIRVEESIYQCCDLAPEARQAIKSLTERLYIGGPLTNSKGQCGYRRCRA 2700
Qy 2701 SGVLTTCGNTLTCTYLKATAACRAAKLODCTMLVNGDDLVI CESAGTOEDAAALRAETE 2760
Db 2701 SGVLTTCGNTLTCTYLKASAAACRAAKLODCTMLVNGDDLVI CESAGTOEDAAALRAETE 2760
Qy 2761 AMTRYSAAPPDPPOPEYDLELITSCSSNVSAHDASGKRVYVLTTRDPTTPLARAAMETAR 2820
Db 2761 AMTRYSAAPPDPPOPEYDLELITSCSSNVSAHDASGKRVYVLTTRDPTTPLARAAMETAR 2820
Qy 2821 HTPNINSLGNITMPTLWARMILMTHFSSILLAQOLEKALDCQIYGACYSIEPLDLPQ 2880
Db 2821 HTPNINSLGNITMPTLWARMILMTHFSSILLAQOLEKALDCQIYGACYSIEPLDLPQ 2880
Qy 2881 IIERLHGLSAPTLHSYSGEINRVASCLURKLGVPPLTRWRHARSVRAKLLSQGGRAATC 2940
Db 2881 IIERLHGLSAPTLHSYSGEINRVASCLURKLGVPPLTRWRHARSVRARLLSQGGRAATC 2940
Qy 2941 GRYLFNNAVTRKLTPTIPAASOLDLSCGFVAGYSGGDIYHLSLRARPWRPPLCILLLSV 3000
Db 2941 GRYLFNNAVTRKLTPTIPAASOLDLSCGFVAGYSGGDIYHLSLRARPWRPPLCILLLSV 3000
Qy 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010

RESULT 11

US-08-904-686A-14
; Sequence 14, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS: CDNA AND ANTIGEN POLYPEPTIDE
; ADDRESSER: Naughton
; ADDRESSER: Armstrong, Westerman, Hattori, McLeland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: McLealand, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3010 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
us-08-904-686A-14

Query Match 95.3%; Score 15253; DB 1; Length 3010;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2840; Conservative 78; Mismatches 92; Indels 0; Gaps 0;

Qy	1	MSTNPKPQRTKRNTRRNPQDKVPGGQIVGGVYLLPRRGPRLGVRATRKASRSQPRG	60
Db	1	MSTNPKPQRTKRNTRRNPQDKVPGGQIVGGVYLLPRRGPRLGVRAPRKTSSRSQPRG	60
Qy	61	RRQIPKARREGRGAWQPGYPWLYGNEGILGWAGWLLSPRGSPSGWGTDPDRRSRLG	120
Db	61	RRQIPKARREGRGAWQPGYPWLYGNEGILGWAGWLLSPRGSPSGWGTDPDRRSRLG	120
Qy	121	KVIDTLTCGFADLMGYIPLVGAFLGGAARALAHGVRVLEDCGVNATGNLPCSFIFLLA	180
Db	121	KVIDTLTCGFADLMGYIPLVGAFLGGAARALAHGVRVLEDCGVNATGNLPCSFIFLLA	180
Qy	181	LLSLCITIPASAYEVNRVNSGIYHVNTDCSNSSIVVEAADVIMHTPGCVPCVOEGNSSRCW	240
Db	181	LLSLCITIPASAYEVNRVNSGIYHVNTDCSNASIVVEAADLIMHTPGCVPCVREGNSSRCW	240
Qy	241	ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYGDLCGSIPLVSQLTFFSPRRHET	300
Db	241	ALTPTLAARNVITPTTIRRHVDLLVGAFAFCSAMVYGDLCGSFVLSQLTFFSPRRHVT	300
Qy	301	VQDCNCSYRGHVSHRNMWMMNWSPTTALVVSOLLRIQAVVDVWAGAHGVLACLA	360
Db	301	LQDCNCSYIPGHVSHRNMWMMNWSPTTALVVSOLLRIQAVVDVWAGAHGVLACLA	360
Qy	361	YYSWGNWAKVLIIVALLFAGVDGTHTTGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW	420
Db	361	YYSWAGNWKVLIIVMLLFAGVDGTHVTGGQAQKTTNRLVSNFASGPKQIQLINTNGSW	420
Qy	421	HINRTALNCNDSLOTGFFAALFYAHKFNSSGCGPERMASCRPIDWFAQGWGPITTKPNS	480
Db	421	HINRTALNCNDSLOTGFLAALFYTHSFNSSGCGPERMAQCRIIDKFDQCGWGPITYAESRS	480

Qy	481	DORPYCWHYAPRPGCVVVPASQVCGPVYCFPTPSPVVGTTDRSGVPTYSWGENETDVMNLN	540
Db	481	DORPYCWHYPPQCTIVPASEVCGPVYCFPTPSPVVGTTDRFGVPTYRWGENETDVLNLN	540
Qy	541	NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNCIGGVGNRTLCPDTCFRKHPKPEATYKCGSG	600
Db	541	NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNCIGGVGNNTLCPDTCFRKHPKPEATYKCGSG	600
Qy	601	PWLTPRCLVDYPIRLMHPYVCTLNFSIPKVRMYVGGVEHRLNAAACNMTWTRGRRCNLEDRRS	660
Db	601	PWLTPRCMVDYPIVRLMHPYVCTVNFITPKVRMYVGGVEHRLNAAACNMTWTRGRRCOLEDRDP	660
Qy	661	ELSPLLSTTEWQILPCAFTTLPALSTGLTHLHQNIVDVQYLXVYGVGSFAFSFAIKWXYIL	720
Db	661	ELSPLLSTTEWQVLPSCFTTLPALSTGLTHLHQNIVDVQYLXVYGVGSFAFSFAIKWXYVL	720
Qy	721	LLFLLADARVCACLWMMLLIAQEALENLVNLNAAASVAGAHGILSFLVFFCAAWYIKG	780
Db	721	LLFLLADARVCACLWMMLLIAQEALENLVNLNAAASVAGAHGILSFLVFFCAAWYIKG	780
Qy	781	RLAPGAAYFYGVWPLLLLLALPPRAYALDREMAASCGGAVLVGLVLTSLSPYKVFILT	840
Db	781	RLVPGATYALYGVWPLLLLLALPPRAYAMDREMAASCGGAVFVGLVLLTSLSPYKVFELA	840
Qy	841	RLIWLQYFITRAEAMQVWVPLNVRGGRDAIILLTCAVHPELIPDITKLLAILGLPLM	900
Db	841	RLIWLQYFTRAADLHWIPLNARGGRDAIILLCAVHPELIPDITKLLAILGLPLM	900
Qy	901	VLOAGITRVPYFVRAOGLIACMLVRKAGHYVQVFMKLGALTGYVNVNHLTPLRDWA	960
Db	901	VLOAGITRVPYFVRAOGLIHACMLVRKAGHYVQVMAFMKLGALTGYIYNHLLTPLRDWP	960
Qy	961	HAGRLDLAVAVEPVVFSAMETKVIITWGADTAAACDIIILGLPVSARRGKEIFLGPADSLEG	1020
Db	961	RAGRLDLAVAVEPVVFSAMETKVIITWGADTAAACDIIILGLPVSARRGKEIFLGPADSLEG	1020
Qy	1021	QGWRLAPITAYSQOQTRGVLCIITSLTGRDKQVEGEVQVVSSTATQSFATCINGVCWT	1080
Db	1021	RGLRLAPITAYSQOQTRGLLCCIITSLTGRDKQVEGEVQVVSSTATQSFATCINGVCWT	1080
Qy	1081	VYHGAGSKTLAGKGPITQMYTNVDLGVQAPPGARSMTPCSCGSDLLVLTTRHADVI	1140
Db	1081	VYHGAGSKTIAAPKGPITQMYTNVDQLVGPKPPGARSMTPCSCGSDLLVLTTRHADVI	1140
Qy	1141	PVRRRGDSRGLSPRPVSYLKGSSGGPLLCPSGHVVGVPRAAVCTRGVAKAVDFIPVES	1200
Db	1141	PVRRRGDSRGLSPRPVSYLKGSSGGPLLCPPGHAVGIFRAAVCTRGVAKAVDFIPVES	1200
Qy	1201	METTMRSPVFTDSTPPPAVPQTFOVAHLHAPTGSKSTKVPAAVAAQGYKVLVLPNSVAA	1260
Db	1201	METTMRSPVFTDSTPPPAVPQSFOVAHLHAPTGSKSTKVPAAVAAQGYKVLVLPNSVAA	1260
Qy	1261	TLGFGAYMSKAHGDIPNIRGTITTTGSGIITYTYKFLADGGCGSGAYDIIICDECHS	1320
Db	1261	TLGFGAYMSKAHGDIPNIRGTITTTGAPVITYTYKFLADGGCGSGAYDIIICDECHS	1320
Qy	1321	TDSTTILGICTVLDOAETAGARLVVLTATATPPGSVTVPHNIBELGLSNNGEIPFYGKAI	1380
Db	1321	TDSTTILGICTVLDOAETAGARLVVLTATATPPGSVTVPHNIBELVNTNGEIPFYGKAI	1380
Qy	1381	PIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAAVAYRGDLDSVIPPIGDWWVATDAL	1440
Db	1381	PIEAIKGRHLIFCHSKKKCDELAALKLTGLGLNAAVAYRGDLDSVIPPIGDWWVATDAL	1440
Qy	1441	MTGFTGDFSDVIDCNTCTVTQVDFSLDPTFTIETTVPDQDAVSRSQRRGTRGRSGIYR	1500
Db	1441	MTGFTGDFSDVIDCNTCTVTQVDFSLDPTFTIETTVPDQDAVSRSQRRGTRGRSGIYR	1500
Qy	1501	FVTGEPSPGMFDSVILCECYDAGCAWVELTPAETSURLRAYLNTPGLPVCQDHLFEFWS	1560
Db	1501	FVTGEPSPGMFDSVILCECYDAGCAWVELTPAETSURLRAYLNTPGLPVCQDHLFEFWS	1560

1561 VFTGLTHIDAFLSQTQAGDNFPYLVAIQATVCARAQAPPSWDQMKCLIRLKPTLHG 1620
1561 VFTGLTHIDAFLSQTQAGDNFPYLVAIQATVCARAQAPPSWDQMKCLIRLKPTLHG 1620
1621 PTPLLYRIGAVONEVILTHPTIKYIMACHSADLEVVTTWLVGVGLAALAAAYCLTTGSV 1680
1621 PTPLLYRIGAVONEVILTHPTIKYIMACHSADLEVVTTWLVGVGLAALAAAYCLTTGSV 1680
1681 VTVGRIILSGKPAVVDREVLVQBEDEMEECASQLPYIEOGMOLAEQFKOKALGELLQAT 1740
1681 VTVGRIILSGRPAIVPDRELLVQBEDEMEECASHLPYIEOGMOLAEQFKOKALGELLQAT 1740
1741 KOAEAAAPVSEKRALETFWAKHWNFTISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
1741 KOAEAAAPVSEKRALETFWAKHWNFTISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
1801 LTTQNTLLFNILGGWVAQAAPPSAASAPVAGIAGAAGVSGIGLKVLDIILAGYGAGVA 1860
1801 LTTQNTLLFNILGGWVAQAAPPSAASAPVAGIAGAAGVSGIGLKVLDIILAGYGAGVA 1860
1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPCEGAVQWNNRLI 1920
1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPCEGAVQWNNRLI 1920
1921 AFASRGNHVSPTHYVPESDAAARVQTQILSSLTITQLLKRLHQMINEDCSTPCSGSMLRDV 1980
1921 AFASRGNHVSPTHYVPESDAAARVQTQILSSLTITQLLKRLHQMINEDCSTPCSGSMLRDV 1980
1981 WDWICTVLTDFTKWLQSKLLPBLPGVFFLSQCGYKGVWVRGDMQTTCCGQAQIAGHVK 2040
1981 WDWICTVLTDFTKWLQSKLLPBLPGVFFLSQCGYKGVWVRGDMQTTCCGQAQIAGHVK 2040
2041 NCSMRIVGRTCSNTHGTFPINAITYTGCTSPAPNYSRALWRVAABEYVEVTVGDPH 2100
2041 NCSMRIVGRTCSNTHGTFPINAITYTGCTSPAPNYSRALWRVAABEYVEVTVGDPH 2100
2101 VYTGMTDNDVCKPCQVPAPPEFTVDGVRHLRYAPACKPLLEDVTFVQGLNQYLVGSOL 2160
2101 VYTGMTDNDVCKPCQVPAPPEFTVDGVRHLRYAPACKPLLEDVTFVQGLNQYLVGSOL 2160
2161 PCEPEPDVTLTSMLTDPGSHITAETAKRLARGSPPSLASSASQLSAPLSKATCTTHHD 2220
2161 PCEPEPDVTLTSMLTDPGSHITAETAKRLARGSPPSLASSASQLSAPLSKATCTTHHD 2220
2221 SPADLIEANLWRQMGNI TRVESENKVILDSPEPLHAEGDEREISVAEILRKSXK 2280
2221 SPADLIEANLWRQMGNI TRVESENKVILDSPEPLHAEGDEREISVAEILRKSXK 2280
2281 FPSALPIWARPDVNPPLLESWKDPDVPVPPVHGCPPLPTKAPPIDPPRRKRTVWLTESNV 2340
2281 FPSALPIWARPDVNPPLLESWKDPDVPVPPVHGCPPLPTKAPPIDPPRRKRTVWLTESNV 2340
2341 SSALAEALATKTCGSSGSSAVDSGTATLPDLASDDGDKGSDVESYSSMPPEGECPDPL 2400
2341 SSALAEALATKTCGSSGSSAVDSGTATLPDLASDDGDKGSDVESYSSMPPEGECPDPL 2400
2401 SDGWSVTSEASEDVEDVCCSMSTWTGALITPCAAEESKLPIPLNSLLRHHNMVYATT 2460
2401 SDGWSVTSEASEDVEDVCCSMSTWTGALITPCAAEESKLPIPLNSLLRHHNMVYATT 2460
2461 SRSASLQKQVTFDRLOVLDHVDYLVKEMAKASTVKAKLSIEBACKLTPPHSAKSKF 2520
2461 SRSASLQKQVTFDRLOVLDHVDYLVKEMAKASTVKAKLSIEBACKLTPPHSAKSKF 2520
2521 GYGAKDVRNLSRAVNHRSVMEDLLEDTETPIDTTIMAKSEVFCVQPEKGGKRPALIV 2580
2521 GYGAKDVRNLSRAVNHRSVMEDLLEDTETPIDTTIMAKSEVFCVQPEKGGKRPALIV 2580
2581 FPDLGVRVCEKMALYDVSTLPOAVMGSSYGFQYSPQORVEFLVNTWKSXKCPMGFSYDT 2640
2581 FPDLGVRVCEKMALYDVSTLPOAVMGSSYGFQYSPQORVEFLVNTWKSXKCPMGFSYDT 2640
2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRARA 2700

2641 RCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRARA 2700
2701 SGVLTTSNGNTLTCYLKATACRAAKLQDCTMLVNGDDLVIICBSAGTQEDAAAALRAFTE 2760
2701 SGVLTTSNGNTLTCYLKATACRAAKLQDCTMLVNGDDLVIICBSAGTQEDAAAALRAFTE 2760
2761 AMTRYSAPPGDPPEYDLELITSCSNVSVHAHDASGKRYYLTRDPTTPLARAAMETAR 2820
2761 AMTRYSAPPGDPPEYDLELITSCSNVSVHAHDASGKRYYLTRDPTTPLARAAMETAR 2820
2821 HTPNSWLGNTIMVAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQ 2880
2821 HTPNSWLGNTIMVAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQ 2880
2881 IIERLHGLSATLTHSYSGEINRVASCLURKLGVPPLRTWRHRARSVRARLLSQGGAATC 2940
2881 IIERLHGLSATLTHSYSGEINRVASCLURKLGVPPLRTWRHRARSVRARLLSQGGAATC 2940
2941 GRYLFNMAVRTKLKLTIPAPASQDLSCGFVAGYSGGDIYHLSLRARPRWFFCLLLLSV 3000
2941 GRYLFNMAVRTKLKLTIPAPASQDLSCGFVAGYSGGDIYHLSLRARPRWFFCLLLLSV 3000
3001 GVGIYLLPNR 3010
3001 GVGIYLLPNR 3010
RESULT 12
US-09-315-850-2
; Sequence 2, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706

/ FILING DATE: 30-JUL-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/769,996
/ FILING DATE: 02-OCT-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/635,451
/ FILING DATE: 28-DEC-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McLeeland, Le-Nhung
/ REGISTRATION NUMBER: 31,541
/ REFERENCE/DOCKET NUMBER: 900703G
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 659-2930
/ TELEFAX: (202) 887-0357
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3010 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-315-850-2

Query Match 95.3%; Score 15253; DB 2; Length 3010;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2840; Conservative 78; Mismatches 92; Indels 0; Gaps 0;

QY	1	MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRKASERSQPRG	60
DB	1	MSTNPKPQRTKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRAPRKTSERSQPRG	60
QY	61	RRQIPKARPEGRWAPQGPWPPLYGNEGILGWAGLLSPRGSPPSGPTDPRRRSRNLG	120
DB	61	RRQIPKARPEGRWAPQGPWPPLYGNEGILGWAGLLSPRGSPPSGPTDPRRRSRNLG	120
QY	121	KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGYNVATGNLPGCSFSIFLLA	180
DB	121	KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGYNVATGNLPGCSFSIFLLA	180
QY	181	LLSCLTTPAGAYEVNRVSGIYHVNTDCSNSIVVEAADVIMHTPGCVPCVOEGNSSRCWV	240
DB	181	LLSCLTTPAGAYEVNRVSGIYHVNTDCSNSIVVEAADVIMHTPGCVPCVREGNSSRCWV	240
QY	241	ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYCDLCGSIPLVLSQLPFTSPRRHET	300
DB	241	ALPTTLAARNVPTIPTTIRRHVDLLVGAFAFCSAMVYGDLCGSFVLVSQLPFTSPRRHVT	300
QY	301	VDQNCSTYPGHVSGHRMADMMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLAGLA	360
DB	301	LQDCNCSYIPGHVSGHRMADMMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLAGLA	360
QY	361	YYSVMGNWAKVLIIVALLFAGVDGETHTTGRVAGHTTSQFTSLFSSGASQKIQLVNTNGSW	420
DB	361	YYSMAGNWKVLIIVMLLFAGVDGTHVTGGQAQKTNLVSMPASGSPQKIQLVNTNGSW	420
QY	421	HINRTALNCNDSLOTGFFAALFYAHKNSSGCPERMASCRPIDMFAQGWGPITYTKPNSS	480
DB	421	HINRTALNCNDSLOTGFLAALFYTHSFNSGCPERMAQCRTIDKFDQGWGPITYVAESRS	480
QY	481	DQPYCHYAPRPGCVVPASQVCGVYCFTPSPVVVGHITDRSGVPTYSWGENETDVMLLN	540
DB	481	DQPYCHYPPQCTIVPASEVCGVYCFTPSPVVVGHITDRFGVPTYSWGENETDVMLLN	540
QY	541	NTRPPQGNWFCGTWNNSGFTKTCGGPPCNTGGVGNRTLIPTDCFRKHPEATYTKCGSG	600
DB	541	NTRPPQGNWFCGTWNNSGFTKTCGGPPCNTGGVGNNTLTCTDCFRKHPEATYTKCGSG	600
QY	601	PWLTPRCLVDYPYRLWHYPCITLNFISIFKVRMYVGGVEHRLNAACNWTGERCNLEDRRS	660
DB	601	PWLTPRCMVDYPYRLWHYPCITVNFIFKVRMYVGGVEHRLNAACNWTGERCDLEDRDP	660
QY	661	ELSPLLSTTTEWQILPCAFPTTLPALSTGLIHLHQNIVDVQVLYGVGSFAVFAIKWEYIL	720
DB	661	ELSPLLSTTTEWQVLPCTFTTLPALSTGLIHLHQNIVDVQVLYGIGSAVSVFAIKWEYVL	720

QY	721	LJFLLLADARVCACLWMMLLIAQEAALLENLVVNAASVAGAHGILSLFLVFFCAAWYIKG	780
DB	721	LJFLLLADARVCACLWMMLLIAQEAALLENLVVNLNSASVAGAHGILSLFLVFFCAAWYIKG	780
QY	781	RLAPGAAYAFYGVWPIILLALLPPRAYALDREMAASCGGAVLVGLVFLTSPYKVFELT	840
DB	781	RLVPGATYALYGVWPIILLALLPPRAYANDREMAASCGGAVFVGLVLLTSPYKVFELA	840
QY	841	RLIWLQYFITRAEAHMVQWVPPNVRGGRDAIILLITCAVHPBELIFDITKLLAILGPLM	900
DB	841	RLIWLQYFTTRAEDLHVWIPPLNARGGRDAIILLMCAVHPBELIFDITKLLAILGPLM	900
QY	901	VLQAGITRVPIYFVRAOGLIRACMLVRKVAGGHVQVMFKLGALTGTYYVNHITPLRDWA	960
DB	901	VLQAGITRVPIYFVRAOGLIHACMLVRKVAGGHVQVMFKLGALTGTYYIYNHITPLRDWP	960
QY	961	HAGIRDLAVAVEPVFSAMETKVIITWGAADTAACGDIILGLPVSARRKEKIFLGPADSLEG	1020
DB	961	RAGIRDLAVAVEPVFSMDMETKIITWGAADTAACGDIILGLPVSARRKEKILLGPAADSLEG	1020
QY	1021	QGNRLLAIPITAYSQOTRGVLGCIITSLTGRDKNQEVEQVQVSTATQSFATCINGVCWT	1080
DB	1021	RGURLLAIPITAYSQOTRGLGCIITSLTGRDKNQEVEQVQVSTATQSFATCINGVCWT	1080
QY	1081	VYHGAGSKTLAGPKGPIITQMYTNVDLVLGVQAPPGARSMTPCSCSSDLYLVTRHADVI	1140
DB	1081	VYHGAGSKTLAAAPKGPITQMYTNVDQDLVGPKPARGSLTPTCTCGSSDLYLVTRHADVI	1140
QY	1141	PVRRRGDSRGLSLSPRPVSYLKGSSGGPLICPSGHVVGVPRAAVCTRGVAKAVDPIPVES	1200
DB	1141	PVRRRGDSRGLSLSPRPVSYLKGSSGGPLICPFHAGVIGIFRAAVCTRGVAKAVDFPVES	1200
QY	1201	MEITMRSPTVDNSTPAPVPTQVAHLHAPTSGSKSTKVPAAVAAOGYKVLNPSVAA	1260
DB	1201	MEITMRSPTVDNSTSPAPVPSQVAHLHAPTSGSKSTKVPAAVAAOGYKVLNPSVAA	1260
QY	1261	TLGFGAYMSKAHGDIPNIRTVRTITGGSTYTYGKFLADGGCSGAYDIIICDECHS	1320
DB	1261	TLGFGAYMSKAHGLDIPNIRTVRTITGGSTYTYGKFLADGGCSGAYDIIICDECHS	1320
QY	1321	TDSTTILIGITVLDQAEATAGARLVVLTATATPGSVTVPHPNIEIIGLSNNGEIPFYGKAI	1380
DB	1321	TDSTTILIGITVLDQAEATAGARLVVLTATATPGSVTVPHPNIEEVALSNTGEIPFYGKAI	1380
QY	1381	PIEAKGGRHLIFCHSKKCDLAAKLTGGLNNAVAYRGLDVSIVPIPIGVVVTATDAL	1440
DB	1381	PIBAIRGGRHLIFCHSKKCDLAAKLSGLGINAVAYRGLDVSIVPTIGDWWVTATDAL	1440
QY	1441	MTGETGDFSDVIDCNTCTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGTRGRSGIYR	1500
DB	1441	MTGYTGDFSDVIDCNTCTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGTRGRSGIYR	1500
QY	1501	FVTPGERPSGMFDSVLCCEYDAGCAWELTPAETSRLRAYLNTPGVPCQDHLFEWES	1560
DB	1501	FVTPGERPSGMFDSVLCCEYDAGCAWELTPAETSRLRAYLNTPGVPCQDHLFEWES	1560
QY	1561	VFTGLTHIDAHFLTSQTKOAGDNPPYLVAQVATVCARAQAPPSPDWQMKCLIRLKPTLHG	1620
DB	1561	VFTGLTHIDAHFLTSQTKOAGDNPPYLVAQVATVCARAQAPPSPDWQMKCLIRLKPTLHG	1620
QY	1621	PTPLLYRLGAVQNEVILTHPTIKYIMACMSADLEVTSTWVLVGVLAALAAAYCLTTGSV	1680
DB	1621	PTPLLYRLGAVQNEVILTHPTIKYIMACMSADLEVTSTWVLVGVLAALAAAYCLTTGSV	1680
QY	1681	VIVGRIILSGKPAVVPDREVLVYQSFDEMEECASQIPIYEQGQLAEOQKALGLLOAT	1740
DB	1681	VIVGRIILSGRPAVVPDRELLVYQSFDEMEECASHLPYIEQGMQLAEOQKALGLLOAT	1740
QY	1741	KQAAAAAVVESKRALETFWAKHWNFIISGIVYLAGLSTLPGNPAIASLMAFTASITSP	1800
DB	1741	KQAAAAAVVESKRALETFWAKHWNFIISGIVYLAGLSTLPGNPAIASLMAFTASITSP	1800


```
Qy 2041 NGSMRIVGPRITCSNTHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDHF 2100
Db 2041 NGSMRIVGKTCNSNTHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDHF 2100
Qy 2101 YVTGMTNTNVKPCQVPAPPEFTEVDGVRLLHRYAPACPLREDVTFQVGLNQYLVGSOL 2160
Db 2101 YVTGMTNTNVKPCQVPAPPEFSEVDGVRLLHRYAPACPLREEVTFQVGLNQYLVGSOL 2160
Qy 2161 PCEPEPDVTVLTSMLTDPESHITAETAKERLARGSPPSLASSASOLSAPSLKACTTHHD 2220
Db 2161 PCEPEPDVAVLTVLSMTDPESHITAETAKRRLARGSPPSLASSASOLSAPSLKACTTHHV 2220
Qy 2221 SPDADLIEANLLWRQEMGNITRVESENKVVILDSFEPLHAEGDREISVAAILRKSRK 2280
Db 2221 SPDADLIEANLLWRQEMGNITRVESENKVVILDSFDPLRAEEDREVSVAAILRKSKK 2280
Qy 2281 FPSALPIWARPDYNPPLLESWKDPDYPPVHVGCPLPPTKAPPIPPRRKRKTVVLTESNV 2340
Db 2281 FPAAMPIWARPDYNPPLLESWKDPDYPPVHVGCPLPPIKAPPIPPRRKRKTVVLTESV 2340
Qy 2341 SSALAEATKTFGSSGSAVDGATATLPDLASDDGDKGSDVESYSSMPLEGEGRDPL 2400
Db 2341 SSALAEATKTFGSSSAVDGATATLPDQASDDGDKGSDVESYSSMPLEGEGRDPL 2400
Qy 2401 SDGSMSTVSEEBASEDVCCSMSTWTGALITPCAABESKLPINPLNSLLRHHNNVYATT 2460
Db 2401 SDGSMSTVSEEBASEDVCCSMSTWTGALITPCAABESKLPINALNSLLRHHNNVYATT 2460
Qy 2461 SRSASLRQKKVTFDRQLVDLDDHYRDLKEMKAKASTVRAKLLSIBEACKLTPPHSAKSKF 2520
Db 2461 SRSAGLRQKKVTFDRQLVDLDDHYRDLKEMKAKASTVRAKLLSVEEACKLTPPHSAKSKF 2520
Qy 2521 GYGAKDVRNLSRAVNHHSVWEDLLEDTEPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVRNLSKAVNHHSVWKLLEDVTPTDITIMAKNVFVCVQPEKGRKPARLIV 2580
Qy 2581 FPDLGVRVCEKMAKYDVSTLTPQAVMGSSYGFQYSPKORVFLVNTWKSCKPMGFSYDT 2640
Db 2581 FPDLGVRVCEKMAKYDVSTLTPQVVMGSSYGFQYSPGQRVEFLVNTWKSCKNPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRCA 2700
Db 2641 RCFDSTVTENDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRCA 2700
Qy 2701 SGVLITSCGNTLTCYLKATAACRAAKLODCTMLVNGDDLVVICESAGTOEDAAALRAETE 2760
Db 2701 SGVLITSCGNTLTCYLKASAAACRAAKLODCTMLVNGDDLVVICESAGTOEDAAALRAETE 2760
Qy 2761 AMTRYSAAPPDPOPEYDLELITSCSSNVSVADHASKRVYLYLTRDPTPLARAAWETAR 2820
Db 2761 AMTRYSAAPPDPOPEYDLELITSCSSNVSVADHASKRVYLYLTRDPTPLARAAWETAR 2820
Qy 2821 HTPINSWLGNIIMYAPTLWARMLMTHFFSILLAQOEKALKDCQIYGACYSIEPLDLPO 2880
Db 2821 HTPVNSWLGNIIMYAPTLWARMLMTHFFSILLAQOEKALKDCQIYGACYSIEPLDLPO 2880
Qy 2881 IIERHLGLSAFTLSHSPGEINRVASCLRLGVPPPLRTHRRARSVRAKLLSQGGRATC 2940
Db 2881 IIERHLGLSAFSLHSYSPGEINRVASCLRLKLVPPPLRVHRARSVRARLLSQGGRATC 2940
Qy 2941 GRYLFNNAVTKLKLTPIPAASOLDLSGMFVAGYSGGDIYHSLSRARPMFPLCILLLSV 3000
Db 2941 GRYLFNNAVTKLKLTPIPAASRLDLSGMFVAGYSGGDIYHSLSRARPMFPLCILLLSV 3000
Qy 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010
```

RESULT 14
US-10-259-275-40
; Sequence 40, Application US/10259275
; Patent No. 6921634

```
; GENERAL INFORMATION:  
; APPLICANT: Lemon, Stanley M.  
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE  
; FILE REFERENCE: 265.0007 0120  
; CURRENT APPLICATION NUMBER: US/10/259,275  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: US 60/171,909  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: US 09/747,419  
; PRIOR FILING DATE: 2000-12-23  
; PRIOR APPLICATION NUMBER: US 60/325,236  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US 60/338,123  
; PRIOR FILING DATE: 2001-11-13  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 40  
; LENGTH: 2985  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence encoded by nucleotides 2077-11121 of SEQ ID  
; OTHER INFORMATION: NO:39  
US-10-259-275-40
```

```
Query Match 92.5%; Score 14801.5; DB 2; Length 2985;  
Best Local Similarity 91.5%; Pred. No. 0;  
Matches 2764; Conservative 89; Mismatches 122; Indels 45; Gaps 5;  
  
Qy 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRKASRSQPRG 60  
Db 1 MSTNPKPQRTKRNTRRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRKATRSQPRG 60  
  
Qy 61 RRQIPKARRPEGRAMVQPGYPWLYKNEGLCWAGWLLSPRGSPSWGPTDPRRRSRNLG 120  
Db 61 RRQIPKARRPEGRAMVQPGYPWLYKNEGMWAGWLLSPRGSPSWGPTDPRRRSRNLG 120  
  
Qy 121 KVIDTLTCGFDADLMGYIPLVCGAPLGGGAARALAHGVRLVEDGVNYATGNLPCGCSIFLLA 180  
Db 121 KVIDTLTCGLADLMGYIPLVGGPLGGGAARALAHGVRLVEDGVNYATGNLPCGCSIFLLA 180  
  
Qy 181 LLSCLITIPASAYEVRNYSYIYHVTNDCSNSSIVYEADAVIMHTPCVPCVQEGNSSRCW 240  
Db 181 LLSCLITVPASAEVRNAGSVYHVTNDCSNSSIVFEADALIMHTPCVPCVQEGNSSRCW 240  
  
Qy 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCASMYVDGLCGSIFLVSQLFTFSRRHET 300  
Db 241 ALTPTLAARNATITPTTIRHHVDLLVGAAALCSAMYVDGLCGSVFLVSQFLTFSRRHAT 300  
  
Qy 301 VQDCNCSYPCGHVSGHRMADMMNWSPTTALVVSOLLRIPOAVVDMVAGAHGWVLAGLA 360  
Db 301 LQDCNCSYPCGHASGHRMADMMNWSPTTALVVSOLLRIPOAVIDMVAGAHGWVLAGLA 360  
  
Qy 361 YYSVMGWNAKVLIVALLFAGVDGETHHTGRVAGHTTSQFTSLFSSGASQKIQLVNTNGSW 420  
Db 361 YYSWAGNAKVLIVMLLFAGVDGHTLTITGCHAARLTSGFAGLTTPGPSORIQLINTNGSW 420  
  
Qy 421 HINTALNCNDSLOTGFFAALFYAHKFNSSGCPERMASCRPIDWFAQGWGPIITYKPNSS 480  
Db 421 HINRTALNCNDSLOTGFLAALFYAHRFNSSGCPERMASCRSIDKPDQGWGPIITYAEPTKD 480  
  
Qy 481 -DORPYCHYIAPRCGVVPASQVCPVYCFTPSPVVVCTTDRSGVPTYSWGNETDVML 539  
Db 481 PDQRYCHYIPQCGIIVPASQVCPVYCFTPSPVVVGTITDRLGNPTYSWGENDTDVLL 540  
  
Qy 540 NNTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLICPTDPRKHPEATYTKCGS 599  
Db 541 NNTRPPQGNWFGCTWMNSTGFTKTCGAPPCNIGGVGNNTLICPTDPRKHPEATYTKCGS 600  
  
Qy 600 GPWLTPRCLVDYPRYLMHYPCTLNFSIFKVRMYGVGVEHRLNAACNWTGRGNCLEDRDR 659  
Db 601 GPWLTPRCWVDYPRYLMHYPCVTNFSIFKVRMYGVGVEHRLNAACNWTGRGNCLEDRDR 660
```

QY	660	SELSPLLLSTTEWOIILPCAPFTTLPALSTGLIHQHONIVDQYLYGVGSAFVSPFAIKWEYI	719
Db	661	SELSPLLLSTTEWOIILPCAPFTTLPALSTGLIHQHONIVDQYLYGVGSAFVSPFAIKWEYV	720
QY	720	LFLFLLLADARVCACILMMMLLIAQAEALNVLNNAASVAGAHGILSLVFFCAAWYIK	779
Db	721	VLFLFLLLADARVCACILMMMLLIAQAEALNVLNNAASVAGAHGILSLVFFCAAWYIK	780
QY	780	GRIPGAAYAFYGWPLPLLLLLLALPPRAYALDREMAASCGGAVLVGLVFTLSPYYKVFL	839
Db	781	GRIPGAAYAFYGAWPLPLLLLLLALPPRAYALDREMAASCGGAVLVGLVFTLSPYYKVFL	840
QY	840	TRLIWLWOYFITRAEAMQWVWPLNVRGORDAIILITCAVHPELIPDIITKLLAILGPL	899
Db	841	ARLLWLWOYFITRAEAMHVLVWVPLNVRGORDAIILITCAVHPELIPDIITKLLAILGPL	900
QY	900	MVLQAGITRVPYFVRAQGLIRACMLVRKVAGGHVQWFMKLGALCTGYVYNHLTPLRDW	959
Db	901	MVLQAGITRVPYFVRAQGLIRACMLVRKVAGGHVQWFMRLGALCTGYVYNHLTPLRDW	960
QY	960	AHAGLRDLAVAPVPVPSAMETKVIITWGAADTAACGDIILGLPVARSARGKEIFLGPADSLE	1019
Db	961	AHAGLRDLAVAPVPVPSAMETKIIITWGAADTAACGDIILGLPVARSARGREILLGPADSLV	1020
QY	1020	GQWRLLAPITAYSQQTRGVLCIITSLTGRDNQVEGEVOVSTATQSFLLATCINGVCW	1079
Db	1021	-----RDNQVEGEVOVSTATQSFLLATCINGVCW	1050
QY	1080	TVYHAGSKTLAGPKGPIITOMYTNVDLILYQWQAPPGARSKTSCSSDLYLVRHADY	1139
Db	1051	TVYHAGSKTLAGPKGPIAQMNTYNDQDLGWMPAPPGARSLTPTCTGSSDLYLVRHADY	1110
QY	1140	IPVRRRGGSGLLSPRPVSLKSGSGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVE	1199
Db	1111	IPVRRRGGSGLLSPRPVSLKSGSGPLLCPSGCHAVGVFRAAVCTRGVAKAVDFIPVE	1170
QY	1200	SMETTMRSPTFTDNSTPPAVPQTQFAVHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVA	1259
Db	1171	SMETTMRSPTFTDNSTPPAVPQTQFAVHLHAPTSGSKSTRVPAAYAAQGYKVLVLPNSVA	1230
QY	1260	ATLGFAGYMSKAHIGIDNIRTVGTITITGSGITVSTYKGFKLADGGCGGAYDIIICDECH	1319
Db	1231	ATLGFAGYMSKAHIGIDNIRTVGTITITGAPITVSTYKGFKLADGGCGGAYDIIIMCDECH	1290
QY	1320	STDSFTTILGTVLDQAEATAGARLVVLATATPPGSGVTVPHENIEIGLSNNGEIPFYGKA	1379
Db	1291	STDSFTTILGTVLDQAEATAGARLVVLATATPPGSGVTVPHENIEEVALSNTGEIPFYGKA	1350
QY	1380	IPIBAIKGGRHLIFCHSKKKCDELAALKITGLGLNNAVAYRYGLDVSVIPPIGDVVVVATDA	1439
Db	1351	IPIBAIKGGRHLIFCHSKKKCDELAALKSLGINAVAYRYGLDVSIVPTSGDVVVVATDA	1410
QY	1440	LMTGFTGDFSDVDCNTCVTVDFSLDPTFTTITTTVPQDAVRSQRGRTGRSGIY	1499
Db	1411	LMTGFTGDFSDVDCNTCVTVDFSLDPTFTTITTTVPQDAVRSQRGRTGRSGIY	1470
QY	1500	RFVTPGPRPSGMFSSVLCYCDAGCAWYELTPAETSURLRAYLNTPLPVCQDHLFEWE	1559
Db	1471	RFVTPGPRPSGMFSSVLCYCDAGCAWYELTPAETSURLRAYLNTPLPVCQDHLFEWE	1530
QY	1560	SVFTGLTHIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSPGWQDQWKKCLIRLKPTLH	1619
Db	1531	SVFTGLTHIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSPGWQDQWKKCLIRLKPTLH	1590
QY	1620	GPTPLLYRLGAVQNEVILTHPIITKYIMACMSADLEWVSTWVLVGGVLAALAAAYCLITGGS	1679
Db	1591	GPTPLLYRLGAVQNEVILTHPIITKYIMACMSADLEWVSTWVLVGGVLAALAAAYCLITGGS	1650
QY	1680	VVIIVGRILLSKGPVPPPREVLVYQBFDEMECASOLPYIEQGMQLAEQFKOKALGILLQTA	1739
Db	1651	VVIIVGRILLSGRPAVPPPREVLVYREFDEMECASHLPYIEQGMQLAEQFKOKALGILLQTA	1710
QY	1740	TKOAEAAAUVESKWEALETFMAKHMWNFTISGIOYLAGLSTLPGNPAIASLMAFTASITS	1799
Db	1711	TKOAEAAAUVESKWEALETFMAKHMWNFTISGIOYLAGLSTLPGNPAIASLMAFTASITS	1770
QY	1800	PLTTQNTLLFNILGGVAAQAAPPASAASAFVVGAGIAGAAVGSIGLGKVLVDIILAGYAGV	1859
Db	1771	PLTTQNTLLFNILGGVAAQAAPPASAASAFVVGAGIAGAAVGSIGLGKVLVDIILAGYAGV	1830
QY	1860	AGALVAFKMSGVRPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNR	1919
Db	1831	AGALVAFKMSGVRPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNR	1890
QY	1920	IAPASRGNHVSPTHYPYSPESDAAARVTQIILSSSLTITOLLKELHOWINEDCSTPCSGSWLRD	1979
Db	1891	IAPASRGNHVSPTHYPYSPESDAAARVTQIILSSSLTITOLLKELHOWINEDCSTPCSGSWLRD	1950
QY	1980	VNDWICTVLTDFTKWLQSKLLPRLPGVPFLSCQRYGVWVRGDIIMQTTCPCGAQIAGHV	2039
Db	1951	VNDWICTVLTDFTKWLQSKLLPRLPGVPFLSCQRYGVWVRGDIIMHTTCPCGAQIAGHV	2010
QY	2040	KNGSMRIIVGPRTCSTNTHGTFPPINAYTTGCTPSPAPNYSRALWRVAABEYEVTRVGDF	2099
Db	2011	KNGSMRIIVGPRTCSTNTHGTFPPINAYTTGCTPSPAPNYSKALWRVAABEYEVTRVGDF	2070
QY	2100	HYVTGTTDNVCKPCOVPAPEFFTEVDGVRHLRYAPACKPLLEDVTFQVGLNQYLVGSO	2159
Db	2071	HYVTGTTDNVCKPCOVPAPEFFTEVDGVRHLRYAPACKPLLEDVTFQVGLNQYLVGSO	2130
QY	2160	L-----PCEPEPDVTVLTSMLTDPESHITAEAKRLARGSPPSLASSASQLSAPSLKAT	2214
Db	2131	LHRYAPVCKP-----LLRDEVFQVGLNQYLVGSQLARGSPPSLASSASQLSAPSLKAT	2185
QY	2215	CTTHH-----DSPADLIEANLLWRQEMGNITRVESENKVVILDSFPFLHAGDEREISV	2270
Db	2186	CTTHSSVNLDSPPVDLITAAANLLWRQEMGNITRVESENKVVILDSFPFLHAGDEREISV	2245
QY	2271	AAELLRKSRFPALPIWARPDPNPLLESWKDPDVPVPPVHGCPLPPTKAPPIPPPRRK	2330
Db	2246	AAELLRKSKFPAAPIWARPDPNPLLESWKDPDVPVPPVHGCPLPPTKAPPIPPPRRK	2305
QY	2331	RTVVLTESNVSSALAEALATKTFGSSGSSAVDSGTATATLPOLASDDGDKGSDVESYSSMPP	2390
Db	2306	RTVVLTSTVSSVLAELATKTFGSSELSAASDGTATATPDTSDNGKSDAESCSMPP	2365
QY	2391	LEGEPPGDDLSDGSSWSTVSEASESDVVCSSMSYTWTCALITPCAAESKULPINPLSNLL	2450
Db	2366	LEGEPPGDDLSDGSSWSTVSEAGESVVCSSMSYTWTCALITPCAAESKULPINPLSNLL	2425
QY	2451	RHNMVYATTSRSASLRQKVTFDRLOVLDHHDYDLVKEMKAKASTVKAKLLSIEEACKL	2510
Db	2426	RHNMVYATTSRSASLRQKVTFDRLOVLDHHDYDLVKEMKAKASTVKAKLLSIEEACKL	2485
QY	2511	TPPHSAKSGYGAQVKNLSSRAVNHIRSWEDELLEDTEPTIDTTIMAKSEVFCVQPEK	2570
Db	2486	TPPHSAKSGYGAQVKNLSSRAVNHIRSWEDELLEDTEPTIDTTIMAKSEVFCVQPEK	2545
QY	2571	GGRPABLIIVFPDLGVRCVKEMALYDVVSTLPQAVMGSSYGFQYSPKQVBEFLVNTWKS	2630
Db	2546	GGRPABLIIVFPDLGVRCVKEMALYDVVSTLPQAVMGSSYGFQYSPKQVBEFLVNTWKS	2605
QY	2631	KCPMGFSYDTRCFDSTVTESDIRVEESIYQCCDLAPARQAIRSLTERLYTGGPLTNSKG	2690
Db	2606	KCPMGFSYDTRCFDSTVTENDIRVEESIYQCCDLAPARQAIRSLTERLYTGGPLTNSKG	2665
QY	2691	QNCGYRCRAGSVLTTSCGNTLTCLYKATACRAKQDCTMLVNGDDLVCSSAGTQE	2750
Db	2666	QNCGYRCRAGSVLTTSCGNTLTCLYKASACRAKQDCTMLVNGDDLVCSSAGTQE	2725
QY	2751	DAALRAFTAMTRYAPPDPPQPEYDLELITSCSSNVSVAHDAAGKRVYLTTRDPTTP	2810
Db	2726	DAALRAFTAMTRYAPPDPPQPEYDLELITSCSSNVSVAHDAAGKRVYLTTRDPTTP	2785
QY	2811	LARAWEATARTPTINSMLGNIIMYAPTILWARMILMTHFFSILLAQEQLEKALDQCIYGAC	2870

Db 2786 IARAWEATARHTPVNSWLNIIIMVAPTLWARMLMTHFFSILLAQEQEKALDCQIYGAC 2845
QY 2871 YSIEPLDLPOIIEHLGLSFTLHYSYSGEINRVASCLRLKGLGVPPLRWRHARSVRAKL 2930
Db 2846 YSIEPLDLPOIIEHLGLSFTLHYSYSGEINRVASCLRLKGLGVPPLRWRHARSVRAKL 2905
QY 2931 LSQGRATCGRYLFNVAVRTKLTLPPIAASOLDLSCWGFVAGYSGGDIYHLSLRARPRW 2990
Db 2906 LSQGRATCGRYLFNVAVRTKLTLPPIAASOLDLSCWGFVAGYSGGDIYHLSLRARPRW 2965
QY 2991 FPLCLLLLSVGVGIYLLPNR 3010
Db 2966 FMLCLLLLSVGVGIYLLPNR 2985
RESULT 15
US-09-014-416-5
; Sequence 5, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne H.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; TITLE OF INVENTION: USSS THEREOF
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentih Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-014-416-5
Query Match 91.5%; Score 14640.5; DB 2; Length 3011;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2697; Conservative 157; Mismatches 156; Indels 1; Gaps 1;
QY 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVILLPRGPRLGVRATRKASERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVILLPRGPRLGVRATRKASERSQPRG 60
QY 61 RROPTPKARRPGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRRSRNILG 120
Db 61 RROPTPKARRPGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRRSRNILG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGAARALAHGVRLVEDGVNYATGNLPGCSFSIFLIA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGAARALAHGVRLVEDGVNYATGNLPGCSFSIFLIA 180
QY 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVYEADVIMHTPGVCVQEGNSRCWY 240
Db 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVYEADVIMHTPGVCVQEGNSRCWY 240
QY 241 ALTPTLAARNASVPTTTIRRHVDLLVGTAAFCSAMVVGDLCSIFLVSQLFTHSPRHHET 300
Db 241 ALTPTLAARNASVPTTTIRRHVDLLVGTAAFCSAMVVGDLCSIFLVSQLFTHSPRHHET 300
QY 301 VQDCNCSIYPGHVSGHRMAWDMWMSPTTALVVSQLLRIPQAVVDMVAGAHGWVLAGLA 360
Db 301 VQDCNCSIYPGHVSGHRMAWDMWMSPTTALVVSQLLRIPQAVVDMVAGAHGWVLAGLA 360
QY 361 YYSMVGNWAKVILLVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
Db 361 YYSMVGNWAKVILLVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
QY 421 HINRTALNCNDSLQGTFFAALFYAHKFNSGCGPERMASCRPIDWFAQGWGPITYTKPNSS 480

Db 421 HINRTALNCNDSLQGTFFAALFYAHKFNSGCGPERMASCRPIDWFAQGWGPITYTKPNSS 480
QY 481 DORPYCWHYAPRPGCVVPASQVCGPVYCFTPSVVGTDTDSGVPVTSYSGWNETDVMLLN 540
Db 481 DORPYCWHYAPRPGCVVPASQVCGPVYCFTPSVVGTDTDSGVPVTSYSGWNETDVMLLN 540
QY 541 NTRPPOQGNWFCTWMNSTGTFTKTCGGPPCNIGGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Db 541 NTRPPOQGNWFCTWMNSTGTFTKTCGGPPCNIGGVGNRTLICPTDCFRKHPEATYTKCGSG 600
QY 601 PWTLPRLCLVDYPYRLWHYPCITLNFSPFKVRMYVGVGEHRLNAACNWTGERCNLDRDRS 660
Db 601 PWTLPRLCLVDYPYRLWHYPCITLNFSPFKVRMYVGVGEHRLNAACNWTGERCNLDRDRS 660
QY 661 ELSPLLLSTTEWQILPCAFITLPAALSTGLIHLHQNIVDVQYLVGVGSFVAFKWEYIL 720
Db 661 ELSPLLLSTTEWQILPCAFITLPAALSTGLIHLHQNIVDVQYLVGVGSFVAFKWEYIL 720
QY 721 LFLLLADARVCACLWMLLIAQAEALLENVLNAASVAGAHGILSLFVFFCAAWYIKG 780
Db 721 LFLLLADARVCACLWMLLIAQAEALLENVLNAASVAGAHGILSLFVFFCAAWYIKG 780
QY 781 RLAPCAAYAFYGVNPLLLLLALLPPRAYALDREMAASCGGAVLVGLVFLTSLPYKVFELT 840
Db 781 RLAPCAAYAFYGVNPLLLLLALLPPRAYALDREMAASCGGAVLVGLVFLTSLPYKVFELT 840
QY 841 RLIIWQLQYFITRAEAHMQVWVPLNVRGGRDAITLLTCAVHPELIFDITKLLLATLGLPLM 900
Db 841 WCMWMLQYFLTRVEAQLHVWVPLNVRGGRDAIVLLMCVHPTLVDFITKLLLATLGLPLM 900
QY 901 VLOAGITRVPYFVRAOGLIRACMLVRKVAGGHYQVMPKMGALGTGVYVNHILTPLRDWA 960
Db 901 ILQASLLKVPFVRVQGLLRICALARKIAGGHYVQMAIIKLGALGTGVYVNHILTPLRDWA 960
QY 961 HAGLRDLAVAVEPVVFSAMETKVIITWGDATACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HNGLRDLAVAVEPVVFSAMETKVIITWGDATACGDIINGLPVSARRGKEIFLGPADGMS 1020
QY 1021 QGWRLLAPITAYSQOTRGVLCIITSLTRGDKNQEVEGVVSTATQSFATCINGVCWT 1080
Db 1021 KGRWLLAPITAYSQOTRGVLCIITSLTRGDKNQEVEGVVSTATQSFATCINGVCWT 1080
QY 1081 VYHAGSXTLAGPKGPIITQMTNVVDLIVGHOAPPGARSMTPCSGSSDLYLVTRHADVI 1140
Db 1081 VYHAGSXTLAGPKGPIITQMTNVVDLIVGHOAPPGARSMTPCSGSSDLYLVTRHADVI 1140
QY 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVEN 1200
QY 1201 METTNRSPVFTDNSTPPPAVQPTFOVAHLHAPTSKSKTKVPAAYAAQGVKVLVLPNSVAA 1260
Db 1201 LGTTNRSPVFTDNSTPPPAVQPTFOVAHLHAPTSKSKTKVPAAYAAQGVKVLVLPNSVAA 1260
QY 1261 TLGFCAYMSKAGHDIPNIRTGVRTITTCGSIITYSYTKFPLADGCGSGGAYDIIICDECHS 1320
Db 1261 TLGFCAYMSKAGHDIPNIRTGVRTITTCGSIITYSYTKFPLADGCGSGGAYDIIICDECHS 1320
QY 1321 TDSTTILGITVLDQAEATAGARLVVLATATPPGSVTVPHNIEEIGLSNNGSIPFYGKAI 1380
Db 1321 TDSTTILGITVLDQAEATAGARLVVLATATPPGSVTVPHNIEEIGLSNNGSIPFYGKAI 1380
QY 1381 PTEALIKGRHILFCHSKKKCDELAALKTGLGLNAVAYYRGLDVSVIPPIGDVWVATDAL 1440
Db 1381 PTEALIKGRHILFCHSKKKCDELAALKTGLGLNAVAYYRGLDVSVIPPIGDVWVATDAL 1440
QY 1441 MTGFTGDSVSDICNTCVTQTVDFSLDPTFTTETTTVPDAYSRSORRGTGKRGSIYR 1500
Db 1441 MTGFTGDSVSDICNTCVTQTVDFSLDPTFTTETTTVPDAYSRSORRGTGKRGSIYR 1500
QY 1501 FVTPGERPSGMPDSSVLCECYDAGCAWVELTTPAETSRLRAYLNTFGLPVCQDHLFEWS 1560

Db 1501 FVAPGRPSGMPDSSVLCECYDAGCAYELTPAETTVRLRAYMNTPCLPVCQDHLSEWEG 1560
Qy 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAOATVCARAQAPPPSWDQWKKLIBLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAOATVCARAQAPPPSWDQWKKLIBLKPTLHG 1620
Qy 1621 PTLLYRLGAVQNEVILTHPIKYIMACMSADLEVTSTWLVGGVLAALAACTLTGTV 1680
Db 1621 PTLLYRLGAVQNEVILTHPIKYIMACMSADLEVTSTWLVGGVLAALAACTLTGTV 1680
Qy 1681 VIIGRIILSGKPAVVPDREVLVYQBFDEMEECASOLPYIEQGMOLAEQFKKALGLLQAT 1740
Db 1681 VIIGRIILSGKPAVVPDREVLVYQBFDEMEECASOLPYIEQGMOLAEQFKKALGLLQAT 1740
Qy 1741 KQAEAAAPVYVESKWRALLETFWAKHMNFISGIIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KQAEAAAPVYVESKWRALLETFWAKHMNFISGIIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Qy 1801 LTTQNTLLFNILGGWAAQALAPSAASAFVAGAGTAGRAVSGIIGLGVLDVILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGGWAAQALAPSAASAFVAGAGTAGRAVSGIIGLGVLDVILAGYGAGVA 1860
Qy 1861 GALVAFKMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEQAVQMMNRLI 1920
Db 1861 GALVAFKMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEQAVQMMNRLI 1920
Qy 1921 AFASRGNHVSPTHVYVPSDAAARVTOILSLITITQLLRLHQMINEPCSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHVYVPSDAAARVTOILSLITITQLLRLHQMINEPCSTPCSGSWLRDV 1980
Qy 1981 WDWICTVLTDFKTLWQSKLPLRGVFPFLSCORGYKGVRGDGMOTTCPCGAQIAGHVK 2040
Db 1981 WDWICTVLTDFKTLWQSKLPLRGVFPFLSCORGYKGVRGDGMOTTCPCGAQIAGHVK 2040
Qy 2041 NGSMRIIVGPRCTNTHGTPINAYTTGCTPPSPAPNVSRALWRVAAEEVVEVTRVGDFF 2100
Db 2041 NGSMRIIVGPRCTNTHGTPINAYTTGCTPPSPAPNVSRALWRVAAEEVVEVTRVGDFF 2100
Qy 2101 YVTGMTTDNVKPCQVPAPEFFTEVDGVRLHRYAPACKPLLEDVTFQVGLNOYLVSQ 2160
Db 2101 YVTGMTTDNVKPCQVPAPEFFTEVDGVRLHRYAPACKPLLEDVTFQVGLNOYLVSQ 2160
Qy 2161 PCEPEPDVLTSMLTDPDSHITATAKRRLARGSPSLASSASQLSAPSKATCTTHHD 2220
Db 2161 PCEPEPDVLTSMLTDPDSHITATAKRRLARGSPSLASSASQLSAPSKATCTTHHD 2220
Qy 2221 SPDAELLEANLLWROEMGNITRVESENKVILDSFPLHAEGDERETISVAAETLRKSRK 2280
Db 2221 SPDAELLEANLLWROEMGNITRVESENKVILDSFPLHAEGDERETISVAAETLRKSRK 2280
Qy 2281 FPSALPIWARPDYNPPLLESKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
Db 2281 FPSALPIWARPDYNPPLLESKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
Qy 2341 SSALAEALATKTFSGSGSANDSGTATAPDLASDDGKGSVSYSSMPPLEGEPDPL 2400
Db 2341 SSALAEALATKTFSGSGSANDSGTATAPDLASDDGKGSVSYSSMPPLEGEPDPL 2400
Qy 2401 SDGWSSTVSEEA-SEDVVCCSMYSWTGALITPCAAESKLPINPLNSLLRHHNMVYAT 2459
Db 2401 SDGWSSTVSEEA-SEDVVCCSMYSWTGALITPCAAESKLPINPLNSLLRHHNMVYAT 2459
Qy 2460 TSRSASLRQKVTDFRLQVLDHVRDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSK 2519
Db 2460 TSRSASLRQKVTDFRLQVLDHVRDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSK 2519
Qy 2520 FGYGAKDVRLNLSRAVNHIRSWEDLLEDTETPTIDTTIMAKSEVFCVQPEKGRKPARLI 2579
Db 2520 FGYGAKDVRLNLSRAVNHIRSWEDLLEDTETPTIDTTIMAKSEVFCVQPEKGRKPARLI 2579
Qy 2580 VFPDLGVRVCBKMALYDVVSTLPQAVMGSSYGFQYSPKQVFEFLVNTWKSKKCPMGFSYD 2639
Db 2580 VFPDLGVRVCBKMALYDVVSTLPQAVMGSSYGFQYSPKQVFEFLVNTWKSKKCPMGFSYD 2639

Qy 2640 TRCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGONCGYRRCR 2699
Db 2641 TRCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGONCGYRRCR 2700
Qy 2700 ASGVLTTSCTGNTLTCVYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFT 2759
Db 2701 ASGVLTTSCTGNTLTCVYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFT 2760
Qy 2760 EAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHADAGKRVVYLTTRDPTTPLARAWEA 2819
Db 2761 EAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHADAGKRVVYLTTRDPTTPLARAWEA 2820
Qy 2820 RHTPINSWLGNIIMYAPTILWARMILMTHPFIILAOBLEKALDCEIYGCYIEPLDLP 2879
Db 2821 RHTPINSWLGNIIMYAPTILWARMILMTHPFIILAOBLEKALDCEIYGCYIEPLDLP 2880
Qy 2880 QITERLHGLSFAFTLHYSYSPGEINRVASCLRLKLGVPPLRTWHRHARSVRAKLLSQGGRAAT 2939
Db 2881 QITERLHGLSFAFTLHYSYSPGEINRVASCLRLKLGVPPLRTWHRHARSVRARLLSRGGRAAI 2940
Qy 2940 CGRYLFNWAVRTKLTPTIPAAASQDLISGMFVAGYSGGDIYHLSRARPRWFPLCLLLLS 2999
Db 2941 CGRYLFNWAVRTKLTPTIPAAASQDLISGMFVAGYSGGDIYHLSRARPRWFPLCLLLLS 3000
Qy 3000 VGVGIYLLPNR 3010
Db 3001 AGVGIYLLPNR 3011

Search completed: December 3, 2005, 08:35:29
Job time : 89 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 03:55:27 ; Search time 3363 Seconds
(without alignments)
19015.083 Million cell updates/sec

Title: US-09-662-454-4
Perfect score: 9595
Sequence: 1 gccagccctgatgggggc.....ggcctctctgcagatcatgt 9595

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9595	100.0	9595	2 AAX24843	Aax24843 Infectiou
2	9595	100.0	9595	4 AAC86939	Aac86939 Nucleotid
3	9595	100.0	9595	12 ADO36222	Ado36222 Hepatitis
4	9595	100.0	9595	12 ADO79396	Ado79396 Hepatitis
5	9593.4	100.0	9595	4 AAF23492	Aaf23492 Infectiou
6	8348.2	87.0	9587	13 ADR82189	Adr82189 Hepatitis
7	8348.2	87.0	9587	13 ADT86646	Adt86646 Hepatitis
8	8339	86.9	9611	13 ADS34713	Ads34713 Hepatitis
9	8311.8	86.6	9413	6 AAD25517	Aad25517 Hepatitis
10	8311.8	86.6	9413	6 AAD49655	Aad49655 Hepatitis
11	8311.8	86.6	9413	10 ADF88596	Adf88596 Hepatitis
12	8310.2	86.6	9413	8 AAL53723	Aal53723 Hepatitis
13	8308.6	86.6	9413	2 AAT03960	Aat03960 Partial H
14	8307	86.6	9413	2 AAQ81559	Aaq81559 Hepatitis
15	8294.2	86.4	9413	2 AAQ80498	Aaq80498 DNA encod
16	8250.2	86.0	9605	6 ABK91411	Abk91411 Hepatitis
17	8248.6	86.0	9605	6 ABK91431	Abk91431 Hepatitis
18	8248.6	86.0	9605	6 ABK91424	Abk91424 Hepatitis
19	8248.6	86.0	9605	6 ABK91429	Abk91429 Hepatitis

20	8248.6	86.0	9605	6	ABK91432	Abk91432 Hepatitis
21	8248.6	86.0	9605	6	ABK91430	Abk91430 Hepatitis
22	8248.6	86.0	9605	6	ABK91428	Abk91428 Hepatitis
23	8248.6	86.0	9605	6	ABK91425	Abk91425 Hepatitis
24	8248.6	86.0	9605	6	ABK91426	Abk91426 Hepatitis
25	8248.6	86.0	9605	6	ABK91433	Abk91433 Hepatitis
26	8247	86.0	9605	6	AAD25332	Aad25332 Hepatitis
27	8237.2	85.8	9608	6	ABK91427	Abk91427 Hepatitis
28	8219.6	85.7	9436	2	AAQ63499	Aaq63499 Blood tra
29	8170.4	85.2	9416	2	AAQ20268	Aaq20268 Non-A, no
30	8170.4	85.2	9416	2	AAQ21829	Aaq21829 Non-A, no
31	8168.8	85.1	9416	2	AAX59394	Aax59394 Non-A, no
32	8137.2	84.8	9400	13	ADR82191	Adr82191 Hepatitis
33	8137.2	84.8	9400	13	ADT86648	Adt86648 Hepatitis
34	8099.8	84.4	9472	2	AAQ33282	Aaq33282 Korean he
35	8067	84.1	9609	6	AAQ33038	Aad33038 HCV-SI fu
36	8055.2	84.0	9414	13	ADR82190	Adr82190 Hepatitis
37	8055.2	84.0	9414	13	ADT86647	Adt86647 Hepatitis
38	8051.4	83.9	9405	2	AAQ40426	Aaq40426 Full-leng
39	8032.8	83.7	9391	2	AAQ64175	Aaq64175 NANBHV E1
40	8021.6	83.6	9391	2	AAQ38959	Aaq38959 Hepatitis
41	7946	82.8	9402	2	AAQ41345	Aaq41345 Human hep
42	7918.6	82.5	11076	3	AAA98965	Aaa98965 Hepatitis
43	7914	82.5	11062	6	AAD25331	Aad25331 Hepatitis
44	7885.8	82.2	9275	14	ABE17072	Aeb17072 Hepatitis
45	7201.4	75.1	9599	2	AAX24833	Aax24833 Infectiou

ALIGNMENTS

RESULT 1
AAX24843
ID AAX24843 standard; DNA; 9595 BP.

XX AAX24843;

DT 21-JUN-1999 (first entry)

DE Infectious hepatitis C virus genotype 1b strain HC-J4 genome.

XX HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
XX assay; antiviral; virucide; ss.

OS Hepatitis C virus.

XX Key Location/Qualifiers
FH CDS 342..9374
FT /*tag= a

XX WO9904008-A2.

XX 28-JAN-1999.

XX 16-JUL-1998; 98WO-US014688.

XX 18-JUL-1997; 97US-0053062P.

XX 27-JAN-1998; 98US-00014416.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Yanagi M, Bukh J, Emerson SU, Purcell RH;

XX WPI; 1999-132252/11.

XX P-PSDB; AAW98022.

XX New isolated hepatitis C virus nucleic acids - used to develop products
XX for the diagnosis, prevention and treatment of HCV infections and for
XX developing screening assays.

XX Claim 3; Fig 14A-F; 126pp; English.

XX The present sequence comprises the nucleic acid sequence of the genome of

QY	1801	GGCAATTACCGGCTCGACCGTGTGTGTCCTACCCCGCTCGAGGTGTGTGTCAGTGT	1860
DB	1801	GGCAATTACCGGCTCGACCGTGTGTGTCCTACCCCGCTCGAGGTGTGTGTCAGTGT	1860
QY	1861	ATTGTTTACCCCAAGCCCTGTGTGTTGGTGGGACACCCGATCGTTCGGGTGCCCTACGT	1920
DB	1861	ATTGTTTACCCCAAGCCCTGTGTGTTGGTGGGACACCCGATCGTTCGGGTGCCCTACGT	1920
QY	1921	ATAGCTGGGGGAGATGACAGACGTGATGCTCTCAACACACGCGTCGSCACAAG	1980
DB	1921	ATAGCTGGGGGAGATGACAGACGTGATGCTCTCAACACACGCGTCGSCACAAG	1980
QY	1981	GCAACTGGTTCCGCTGTACATGATGAATAGTACTGGGTTCACTAAGACGTGCGGAGGTC	2040
DB	1981	GCAACTGGTTCCGCTGTACATGATGAATAGTACTGGGTTCACTAAGACGTGCGGAGGTC	2040
QY	2041	CCCCGTGTAACTCGGGGGGTCGGTAACCGCACTTGATCTGCCCCACGGACTGCTTCC	2100
DB	2041	CCCCGTGTAACTCGGGGGGTCGGTAACCGCACTTGATCTGCCCCACGGACTGCTTCC	2100
QY	2101	GGAGCAACCCGAGGCTACTTACAAAATGTGGCTGGGGCCCTGGTTGACACCTAGGT	2160
DB	2101	GGAGCAACCCGAGGCTACTTACAAAATGTGGCTGGGGCCCTGGTTGACACCTAGGT	2160
QY	2161	GCCTAGTACTACCCATACAGGCTTTGGCACTACCCCTGCACCTCAATTTTCCATCT	2220
DB	2161	GCCTAGTACTACCCATACAGGCTTTGGCACTACCCCTGCACCTCAATTTTCCATCT	2220
QY	2221	TTAAGGTTAGGATGTATGTGGGGGGTGGAGCACAGGCTCAATGCCGATGCAATTGGA	2280
DB	2221	TTAAGGTTAGGATGTATGTGGGGGGTGGAGCACAGGCTCAATGCCGATGCAATTGGA	2280
QY	2281	CTCAGGAGACCGCTGTAACTTGGAGACAGGATAGGTGAGAACTCAGCCCGCTGCTGC	2340
DB	2281	CTCAGGAGACCGCTGTAACTTGGAGACAGGATAGGTGAGAACTCAGCCCGCTGCTGC	2340
QY	2341	TGTCACACAGAGTGCAGATACTGCCCTGTGCTTTCACACCCCTACCGGCTTTATCCA	2400
DB	2341	TGTCACACAGAGTGCAGATACTGCCCTGTGCTTTCACACCCCTACCGGCTTTATCCA	2400
QY	2401	CTGGTTTGATCCATCCATCAGAAATCGTGGACGTGCAATACCTGTACGGTGTAGGCT	2460
DB	2401	CTGGTTTGATCCATCCATCAGAAATCGTGGACGTGCAATACCTGTACGGTGTAGGCT	2460
QY	2461	CAGGTTTGTCTCTTTGCAATCAATGGAGTACATCTGTTGCTTTTCTCTCTCTGG	2520
DB	2461	CAGGTTTGTCTCTTTGCAATCAATGGAGTACATCTGTTGCTTTTCTCTCTCTGG	2520
QY	2521	CAGACGCGCGTGTGCTGCTGTGGATGATGCTGTGATAGCCAGGCTGAGGCGG	2580
DB	2521	CAGACGCGCGTGTGCTGCTGTGGATGATGCTGTGATAGCCAGGCTGAGGCGG	2580
QY	2581	CCTTAGAGAACTTGGTGTCTCAATGGCGGCTCCGTGGCGGAGCGCATGTTCTCT	2640
DB	2581	CCTTAGAGAACTTGGTGTCTCAATGGCGGCTCCGTGGCGGAGCGCATGTTCTCT	2640
QY	2641	CCTTTCTGTGTCTCTGCGCGCTGTGATCAATTAAGGGCAGGCTGGCTCTGGGGGG	2700
DB	2641	CCTTTCTGTGTCTCTGCGCGCTGTGATCAATTAAGGGCAGGCTGGCTCTGGGGGG	2700
QY	2701	CGTATGCTTTTATGGGTATGGCGGTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT	2760
DB	2701	CGTATGCTTTTATGGGTATGGCGGTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT	2760
QY	2761	CTTACGCTTGGACCGGAGATGCTGCATCGTGGGGGGTGGGTTCTTGTAGTCTGG	2820
DB	2761	CTTACGCTTGGACCGGAGATGCTGCATCGTGGGGGGTGGGTTCTTGTAGTCTGG	2820
QY	2821	TATTTCTTGACCTGTACCACTACAAAGTGTCTCTCACTAGGCTCATATGGTGGTTAC	2880
DB	2821	TATTTCTTGACCTGTACCACTACAAAGTGTCTCTCACTAGGCTCATATGGTGGTTAC	2880

QY	2881	AATACCTTTATCACCAGACCGGCGCACATCAAGTGTGGTCCCGCCCTCAAGCTTC	2940
DB	2881	AATACCTTTATCACCAGACCGGCGCGCACATCAAGTGTGGTCCCGCCCTCAAGCTTC	2940
QY	2941	GGGGAGGCGCGCATGCATATCCTCCTCACGTGTGCGGTTCATCCAGAGTTAAATTTTG	3000
DB	2941	GGGGAGGCGCGCATGCATATCCTCCTCACGTGTGCGGTTCATCCAGAGTTAAATTTTG	3000
QY	3001	ACATCACAACAACTCCTGTCTGCCATATCAGGCGCGCTCATGTGTCTCCAGGCTGGCATAA	3060
DB	3001	ACATCACAACAACTCCTGTCTGCCATATCAGGCGCGCTCATGTGTCTCCAGGCTGGCATAA	3060
QY	3061	CGAGAGTCCGCTACTTCTGTGCGGCTCAAGGCTCATTTGTCATGCTGTAGTCCGAA	3120
DB	3061	CGAGAGTCCGCTACTTCTGTGCGGCTCAAGGCTCATTTGTCATGCTGTAGTCCGAA	3120
QY	3121	NAAGTCCCGGGGGTCAATATGTCCAAATGCTTTCAATGAAGCTGGGCGGCTCACAGGTA	3180
DB	3121	NAAGTCCCGGGGGTCAATATGTCCAAATGCTTTCAATGAAGCTGGGCGGCTCACAGGTA	3180
QY	3181	CGTACGTTTATAACCACTTTACCCCACTCGGGGACTGGGCCCAACGCGGGCTTACGAGACC	3240
DB	3181	CGTACGTTTATAACCACTTTACCCCACTCGGGGACTGGGCCCAACGCGGGCTTACGAGACC	3240
QY	3241	TTGCGGTGGCGGTAGAGCCCGTCTCTTCCGCCATGGAGACCAAGGTCAATCACTGGG	3300
DB	3241	TTGCGGTGGCGGTAGAGCCCGTCTCTTCCGCCATGGAGACCAAGGTCAATCACTGGG	3300
QY	3301	GAGACACACCGCTGCTGTGGGGAATCATCTTTGGGTCTACCCGCTCTCCGCCCGAAGGG	3360
DB	3301	GAGACACACCGCTGCTGTGGGGAATCATCTTTGGGTCTACCCGCTCTCCGCCCGAAGGG	3360
QY	3361	GGAGAGATATTTTTGGGACCGGCTAGTAGTCTCGAAGGCAAGGTTGCGGACTCCTTG	3420
DB	3361	GGAGAGATATTTTTGGGACCGGCTAGTAGTCTCGAAGGCAAGGTTGCGGACTCCTTG	3420
QY	3421	CGCCCATACGGCTACTCTCCAAACAAACGCGGGCGTACTTGGTTGATCATCACTAGCC	3480
DB	3421	CGCCCATACGGCTACTCTCCAAACAAACGCGGGCGTACTTGGTTGATCATCACTAGCC	3480
QY	3481	TCACAGCGGGGACAAAGCAACAGGTGGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC	3540
DB	3481	TCACAGCGGGGACAAAGCAACAGGTGGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC	3540
QY	3541	AATCTTTTCTGGGACCTGCATCAACGCGGTGTGTGGACTGTCTACCATGGGCTGGCT	3600
DB	3541	AATCTTTTCTGGGACCTGCATCAACGCGGTGTGTGGACTGTCTACCATGGGCTGGCT	3600
QY	3601	CGAAGACCTTAGCGGTCCAAAAGGTCCAAATCAACCAATGTACCAATGTAGACCTGG	3660
DB	3601	CGAAGACCTTAGCGGTCCAAAAGGTCCAAATCAACCAATGTACCAATGTAGACCTGG	3660
QY	3661	ACCTGTGGCTGGCAGGCGCCCGGGGGCGCTCCATGACACCATGAGCTGTGGCA	3720
DB	3661	ACCTGTGGCTGGCAGGCGCCCGGGGGCGCTCCATGACACCATGAGCTGTGGCA	3720
QY	3721	GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCAATTCGGGTGCGCGCGAGGCG	3780
DB	3721	GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCAATTCGGGTGCGCGCGAGGCG	3780
QY	3781	ACAGCAGGGAAGTCTACTCTCCCGGAGGCGCTCTCTCTACCTGAAAGGCTCTCGGGTG	3840
DB	3781	ACAGCAGGGAAGTCTACTCTCCCGGAGGCGCTCTCTCTACCTGAAAGGCTCTCGGGTG	3840
QY	3841	GTCCATTTGCTTGGCTTCGGGCACTGCTGGGCGTCTTCGGGCTGCTGTGTGACCC	3900
DB	3841	GTCCATTTGCTTGGCTTCGGGCACTGCTGGGCGTCTTCGGGCTGCTGTGTGACCC	3900
QY	3901	GGGGGGTGGCAAGGCGGTGGAATTCATACCCGTTGAGTCTATGGAACTACCATGGGT	3960
DB	3901	GGGGGGTGGCAAGGCGGTGGAATTCATACCCGTTGAGTCTATGGAACTACCATGGGT	3960
QY	3961	CTCCGGTCTTCACAGACAACTCAACCCCGGCTGTACCCGAGACATTCACAGTGGCAC	4020

Db 3961 CTCGGTCTTACAGACAACTCAACCCCGGCTGTACCGCAGACATTCCAAGTGGCAC 4020
QY 4021 ATCTGCA CGCTCTACTGGCAGCGCAAGAGACACAAAGTGC CGGCTCGTATG CAGCCC 4080
Db 4021 ATCTGCA CGCTCTACTGGCAGCGCAAGAGACACAAAGTGC CGGCTCGTATG CAGCCC 4080
QY 4081 AAGGTAC AAGGTGCTCGTCTGTAACCCCGTCCGTTGCCGCCACCTTAGGGTTTGGGGCGT 4140
Db 4081 AAGGTAC AAGGTGCTCGTCTGTAACCCCGTCCGTTGCCGCCACCTTAGGGTTTGGGGCGT 4140
QY 4141 ATATGTCCAAGGSCACACGGTATCGACCTTAACATCAGAACTGGGTAAGGACCAATTACCA 4200
Db 4141 ATATGTCCAAGGSCACACGGTATCGACCTTAACATCAGAACTGGGTAAGGACCAATTACCA 4200
QY 4201 CGGGCGGCTCATTAAGTACTCCACTATGGCAAGTTCCTTTGCCGACGGTGGCTTTCTG 4260
Db 4201 CGGGCGGCTCATTAAGTACTCCACTATGGCAAGTTCCTTTGCCGACGGTGGCTTTCTG 4260
QY 4261 GGGGGCCCTATGACATCATATATGTGATGATGCCACTCAACTGACTCGACTACCATCT 4320
Db 4261 GGGGGCCCTATGACATCATATATGTGATGATGCCACTCAACTGACTCGACTACCATCT 4320
QY 4321 TGGGCATCGGCACACTCTGGACCAAGCGGAGACGGCTGGAGCGCGGCTCGTCTGCTCG 4380
Db 4321 TGGGCATCGGCACACTCTGGACCAAGCGGAGACGGCTGGAGCGCGGCTCGTCTGCTCG 4380
QY 4381 CCACCGCTACACTCCGGGATCGGTAGCTGCCACACCCCAATATCGAGGAATAGGCC 4440
Db 4381 CCACCGCTACACTCCGGGATCGGTAGCTGCCACACCCCAATATCGAGGAATAGGCC 4440
QY 4441 TGTCCAA CAATGGAGAGATCCCTTTCTATGGCAAGCCATCCCATAGGSCCATCAAGG 4500
Db 4441 TGTCCAA CAATGGAGAGATCCCTTTCTATGGCAAGCCATCCCATAGGSCCATCAAGG 4500
QY 4501 GGGGAGGCACTCATTTCTGCCATTCAGAGAAATGACAGCTCGCGGCCAAAGC 4560
Db 4501 GGGGAGGCACTCATTTCTGCCATTCAGAGAAATGACAGCTCGCGGCCAAAGC 4560
QY 4561 TGACAGGCTCGGACTGAACGCTGTAGCATATACCGGGGCTTCATGTGTCGTCATAC 4620
Db 4561 TGACAGGCTCGGACTGAACGCTGTAGCATATACCGGGGCTTCATGTGTCGTCATAC 4620
QY 4621 CGCCTATCGGAGACGTCGTTGTGTGGCAACAGACGCTCTAATGACGGGTTTCAACGGCG 4680
Db 4621 CGCCTATCGGAGACGTCGTTGTGTGGCAACAGACGCTCTAATGACGGGTTTCAACGGCG 4680
QY 4681 ATTTTGACTCAGTGATCGACTGCAATACATGTGTCAACGACAGCTCGACTTCAGCTTGG 4740
Db 4681 ATTTTGACTCAGTGATCGACTGCAATACATGTGTCAACGACAGCTCGACTTCAGCTTGG 4740
QY 4741 ATCCACCTTCACCATTCAGACGACCGTGCCGCCAAGACGCGGTGTCGCTCGCAAC 4800
Db 4741 ATCCACCTTCACCATTCAGACGACCGTGCCGCCAAGACGCGGTGTCGCTCGCAAC 4800
QY 4801 GCGAGGTAGAACTGGCAGGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGAGAAAC 4860
Db 4801 GCGAGGTAGAACTGGCAGGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGAGAAAC 4860
QY 4861 GGCCTCGGGCATGTCGATCTTCGGTCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
Db 4861 GGCCTCGGGCATGTCGATCTTCGGTCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
QY 4921 GGTATGAGCTACGCCCGCTGAGACTCGGTTAGGTTGGGGCTTACCTAAATACACCAG 4980
Db 4921 GGTATGAGCTACGCCCGCTGAGACTCGGTTAGGTTGGGGCTTACCTAAATACACCAG 4980
QY 4981 GGTTCGCCGCTGCGCAGGACCACTTGGAGTTCTGGAGAGCGCTTTCACAGGCTCAACC 5040
Db 4981 GGTTCGCCGCTGCGCAGGACCACTTGGAGTTCTGGAGAGCGCTTTCACAGGCTCAACC 5040
QY 5041 ACATAGATGCCCACTTCCTGTCCAGACTAAACAGGAGGAGCAAACTTCCTTACCTGG 5100

Db 5041 ACATAGATGCCCACTTCCTGTCCAGACTAAACAGGAGGAGACAACTTTCCTTACCTGG 5100
QY 5101 TGGCATATCAAGCTTACAGTGTGCCCGAGGCTCAAGCTCCACCTCCATCGTGGGACCAA 5160
Db 5101 TGGCATATCAAGCTTACAGTGTGCCCGAGGCTCAAGCTCCACCTCCATCGTGGGACCAA 5160
QY 5161 TGTGGAAGTGTCTCATACGGCTGAAACCTACACATGACGGGCCAACCCCTCTGTGTATA 5220
Db 5161 TGTGGAAGTGTCTCATACGGCTGAAACCTACACATGACGGGCCAACCCCTCTGTGTATA 5220
QY 5221 GGCTAGAGAGCGCTCCAAAATGAGGTCACTCTCACACACCCCATAACTAAATACATCATGG 5280
Db 5221 GGCTAGAGAGCGCTCCAAAATGAGGTCACTCTCACACACCCCATAACTAAATACATCATGG 5280
QY 5281 CATGCATGTGGCTGACCTGAGAGTGTCTACTAGCACCTGGGTGTGTAGCGGAGTCC 5340
Db 5281 CATGCATGTGGCTGACCTGAGAGTGTCTACTAGCACCTGGGTGTGTAGCGGAGTCC 5340
QY 5341 TTGCAGCTTTGGCCGCATACCTGCTGACGACGAGCAGTGTGTCATTTGTGGCAGGATCA 5400
Db 5341 TTGCAGCTTTGGCCGCATACCTGCTGACGACGAGCAGTGTGTCATTTGTGGCAGGATCA 5400
QY 5401 TCTTGTCCGGGAAGCCAGCTGTCTGTTCCGACAGGGAAGTCTCTACACAGAGTTTCGATG 5460
Db 5401 TCTTGTCCGGGAAGCCAGCTGTCTGTTCCGACAGGGAAGTCTCTACACAGAGTTTCGATG 5460
QY 5461 AGATGGAGAGTGTGCCCTCAAACTTCTTATCATCGACAGGGAATGAGCTGCGCGAGC 5520
Db 5461 AGATGGAGAGTGTGCCCTCAAACTTCTTATCATCGACGAGGAATGAGCTGCGCGAGC 5520
QY 5521 AATTCAAGCAAAAGCGGCTCGGGTTGTGCAAAAGCCACCAAGCAAGCGGAGGCTGCTG 5580
Db 5521 AATTCAAGCAAAAGCGGCTCGGGTTGTGCAAAAGCCACCAAGCAAGCGGAGGCTGCTG 5580
QY 5581 CTCCTGTGTGAGTCCAAAGTGGCGAGCCCTTTAGACCTCTGGCGAAGCACATGTGGA 5640
Db 5581 CTCCTGTGTGAGTCCAAAGTGGCGAGCCCTTTAGACCTCTGGCGAAGCACATGTGGA 5640
QY 5641 ATTTTCATCAGCGGAATACAGTACCTTAGCAGGCTTATCACTCTGCTGGAACCCCGGGA 5700
Db 5641 ATTTTCATCAGCGGAATACAGTACCTTAGCAGGCTTATCACTCTGCTGGAACCCCGGGA 5700
QY 5701 TAGCATATTGATGGCATTTACAGGTTCTATCATAGCCCCGTACCAACCCAAAACACCC 5760
Db 5701 TAGCATATTGATGGCATTTACAGGTTCTATCATAGCCCCGTACCAACCCAAAACACCC 5760
QY 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAATCGCTCCTCCAGCGCTGGT 5820
Db 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAATCGCTCCTCCAGCGCTGGT 5820
QY 5821 CAGCTTTCTGGGCGCCGGCATCGCGGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGG 5880
Db 5821 CAGCTTTCTGGGCGCCGGCATCGCGGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGG 5880
QY 5881 TGCTCGTGGACATCTTGGCGGCTATGGGGCAGGGGTAGCGGCGACTCGTGGCCTTTA 5940
Db 5881 TGCTCGTGGACATCTTGGCGGCTATGGGGCAGGGGTAGCGGCGACTCGTGGCCTTTA 5940
QY 5941 AGGTCAAGAGCGGAGGTGCCCTCCACGAGGACCTGGTCAACTTACTCCCTGCAATCC 6000
Db 5941 AGGTCAAGAGCGGAGGTGCCCTCCACGAGGACCTGGTCAACTTACTCCCTGCAATCC 6000
QY 6001 TCTCTCTGTGGCTGCTGCTCGGGTGTGGCGAGCAATACTGCTGCGCAGCTGG 6060
Db 6001 TCTCTCTGTGGCTGCTGCTCGGGTGTGGCGAGCAATACTGCTGCGCAGCTGG 6060
QY 6061 GCCCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTCCTTCGCGGGGTA 6120
Db 6061 GCCCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTCCTTCGCGGGGTA 6120
QY 6121 ACCAGCTCTCCCTACGCACTATGTGCTGAGAGGACGCTGACGACGTGTCACTCAGA 6180
Db 6121 ACCAGCTCTCCCTACGCACTATGTGCTGAGAGGACGCTGACGACGTGTCACTCAGA 6180

QY	6181	TCCTCTAGCCTTACCATCACTCACTGCTGAAGCGGCTCCACCAAGTGGATTAATAGG	6240
DB			
QY	6181	TCCTCTAGCCTTACCATCACTCACTGCTGAAGCGGCTCCACCAAGTGGATTAATAGG	6240
DB			
QY	6241	ACTGCTTAGCCATGCTCGGCTCGTGGCTGAAGGATGTTGGATGGATGACCG	6300
DB			
QY	6241	ACTGCTTAGCCATGCTCGGCTCGTGGCTGAAGGATGTTGGATGGATGACCG	6300
DB			
QY	6301	TGTTGACTGACTTCAAGACTGGCTCCAGTCCAAACTCTCTGCCGCGGTTACCGGGAGTCC	6360
DB			
QY	6301	TGTTGACTGACTTCAAGACTGGCTCCAGTCCAAACTCTCTGCCGCGGTTACCGGGAGTCC	6360
DB			
QY	6361	CTTTCTCTGTCATGCCAAACCGGGTACAAGGGAGTCTGGCGGGGGACCGCATCATCAAA	6420
DB			
QY	6361	CTTTCTCTGTCATGCCAAACCGGGTACAAGGGAGTCTGGCGGGGGACCGCATCATCAAA	6420
DB			
QY	6421	CCACTGCCATCGGAGACAGATCGCCGGAATGTCTCAAAAACGGTTCATGAGATCG	6480
DB			
QY	6421	CCACTGCCATCGGAGACAGATCGCCGGAATGTCTCAAAAACGGTTCATGAGATCG	6480
DB			
QY	6481	TAGGGCTTAGAACCTGACGAAACAGTGGCACGAAAGCTTCCCATCAACGCATACACCA	6540
DB			
QY	6481	TAGGGCTTAGAACCTGACGAAACAGTGGCACGAAAGCTTCCCATCAACGCATACACCA	6540
DB			
QY	6541	CGGGACCTTGACACACCTCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGGTGGCTG	6600
DB			
QY	6541	CGGGACCTTGACACACCTCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGGTGGCTG	6600
DB			
QY	6601	CTGAGGAGTACGTTGAGGTTACCGTGTGGGGATTTCCACTAGTGAACGGGATGACCA	6660
DB			
QY	6601	CTGAGGAGTACGTTGAGGTTACCGTGTGGGGATTTCCACTAGTGAACGGGATGACCA	6660
DB			
QY	6661	CTGACAAAGTAAAGTGGCATGCGTTCGGGCGCCGGAATTTCTCACGGAGTGGATG	6720
DB			
QY	6661	CTGACAAAGTAAAGTGGCATGCGTTCGGGCGCCGGAATTTCTCACGGAGTGGATG	6720
DB			
QY	6721	GAGTGGGTTGACAGGTACGCTCCGGCGTCAAACTCTTCTACGGGAGGAGCTCACGT	6780
DB			
QY	6721	GAGTGGGTTGACAGGTACGCTCCGGCGTCAAACTCTTCTACGGGAGGAGCTCACGT	6780
DB			
QY	6781	TCCAGGTCGGGCTCAACCAATACTTGGTCGGGTGCGAGCTCCCATGCGAGCCGAAACCG	6840
DB			
QY	6781	TCCAGGTCGGGCTCAACCAATACTTGGTCGGGTGCGAGCTCCCATGCGAGCCGAAACCG	6840
DB			
QY	6841	ACGTAAACAGTGTCTTACCTCCATGCTCACCGATCCCTCCCACTTACAGAGACGGCTA	6900
DB			
QY	6841	ACGTAAACAGTGTCTTACCTCCATGCTCACCGATCCCTCCCACTTACAGAGACGGCTA	6900
DB			
QY	6901	AGCGTAGGCTGGCTAGAGGGTCTCCCGCTTTTAGCCAGCTCATCAGCTAGCCAGTTGT	6960
DB			
QY	6901	AGCGTAGGCTGGCTAGAGGGTCTCCCGCTTTTAGCCAGCTCATCAGCTAGCCAGTTGT	6960
DB			
QY	6961	CTGGCGCTTCTTTGAAGCGCATGCACTACCCACCATGACTCCCGGAGCGCTGACCTCA	7020
DB			
QY	6961	CTGGCGCTTCTTTGAAGCGCATGCACTACCCACCATGACTCCCGGAGCGCTGACCTCA	7020
DB			
QY	7021	TCGAGGCGCAACCTCTTGTGGCGGAGAGATGGCGGAAACATCATCTCGGTGAGTCA	7080
DB			
QY	7021	TCGAGGCGCAACCTCTTGTGGCGGAGAGATGGCGGAAACATCATCTCGGTGAGTCA	7080
DB			
QY	7081	AGAAAGGTAGTAACTCTGACTCTTTCGAAACCGCTTACGCGGAGGGGATGAGAGGG	7140
DB			
QY	7081	AGAAAGGTAGTAACTCTGACTCTTTCGAAACCGCTTACGCGGAGGGGATGAGAGGG	7140
DB			
QY	7141	AGATATCGTCGCGCGAGATCTTCGAAATCCAGGAAGTCCCTTCAGCGTTGCCCA	7200
DB			
QY	7141	AGATATCGTCGCGCGAGATCTTCGAAATCCAGGAAGTCCCTTCAGCGTTGCCCA	7200
DB			
QY	7201	TATGGCACGCGCGGACTTACAACTCTCCACTGCTAGAGTCTCTGGAAGACCCGAGTACG	7260
DB			
QY	7201	TATGGCACGCGCGGACTTACAACTCTCCACTGCTAGAGTCTCTGGAAGACCCGAGTACG	7260
DB			

QY	7261	TCCCTCCGGTGTACACGGATGCCATTGGCCACCTACCAAGGCTCCTCCAATACCACTTC	7320
DB			
QY	7261	TCCCTCCGGTGTACACGGATGCCATTGGCCACCTACCAAGGCTCCTCCAATACCACTTC	7320
DB			
QY	7321	CACGAGAAAGAGGACGGTGTCTTGACAGAAATCAATGTGTCTTCTGCTTGGCGGAGC	7380
DB			
QY	7321	CACGAGAAAGAGGACGGTGTCTTGACAGAAATCAATGTGTCTTCTGCTTGGCGGAGC	7380
DB			
QY	7381	TGCGCACTAAGACCTTCGGTAGCTCCGATCGTGGCGGTGATAGCGGACCGGACCG	7440
DB			
QY	7381	TGCGCACTAAGACCTTCGGTAGCTCCGATCGTGGCGGTGATAGCGGACCGGACCG	7440
DB			
QY	7441	CCCTTCTGACCTCGCTCCGACGAGTGAACAAAGGATCCGACGTTGAGTCTGCTCTCT	7500
DB			
QY	7441	CCCTTCTGACCTCGCTCCGACGAGTGAACAAAGGATCCGACGTTGAGTCTGCTCTCT	7500
DB			
QY	7501	CCATGCCCCCTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTTA	7560
DB			
QY	7501	CCATGCCCCCTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTTA	7560
DB			
QY	7561	CCGTGAGTGAGGAGGCTAGTGAGGATGCTGTGTGCTCAATGTCTTATATCGTGGACAG	7620
DB			
QY	7561	CCGTGAGTGAGGAGGCTAGTGAGGATGCTGTGTGCTCAATGTCTTATATCGTGGACAG	7620
DB			
QY	7621	GGCCCTGATACGCCATCGCTCGGAGGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG	7680
DB			
QY	7621	GGCCCTGATACGCCATCGCTCGGAGGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG	7680
DB			
QY	7681	ACTTTTGTGCTGCTACCAACATGCTTACGCCAACATGCTTACGCCAACATGCTTACGCC	7740
DB			
QY	7681	ACTTTTGTGCTGCTACCAACATGCTTACGCCAACATGCTTACGCCAACATGCTTACGCC	7740
DB			
QY	7741	GGCAGAAAGGTCACCTTTTGACAGATTGCAAGTCTCGATGATCATTAACCGGAGCTAC	7800
DB			
QY	7741	GGCAGAAAGGTCACCTTTTGACAGATTGCAAGTCTCGATGATCATTAACCGGAGCTAC	7800
DB			
QY	7801	TCAAGAGATGAAGGCGAAGGCTCCACAGTTAAGCTAAGCTTCTATCTATAGAGGAGG	7860
DB			
QY	7801	TCAAGAGATGAAGGCGAAGGCTCCACAGTTAAGCTAAGCTTCTATCTATAGAGAGG	7860
DB			
QY	7861	CTGCAAGCTGACGCCCGCCACATTCGCCCAAAATTCAAAATTTGGCTATGGGCAAGGACG	7920
DB			
QY	7861	CTGCAAGCTGACGCCCGCCACATTCGCCCAAAATTCAAAATTTGGCTATGGGCAAGGACG	7920
DB			
QY	7921	TCCGGAACCTTATCCAGCAGCGCGTTAAACACATCCGCTCCGTGCTGGAGGACTTGTGG	7980
DB			
QY	7921	TCCGGAACCTTATCCAGCAGCGCGTTAAACACATCCGCTCCGTGCTGGAGGACTTGTGG	7980
DB			
QY	7981	AAGACACTGAAAACCAATTTGACACCACTATCATGCAAAAGTGAAGTCTTCTGCGTCC	8040
DB			
QY	7981	AAGACACTGAAAACCAATTTGACACCACTATCATGCAAAAGTGAAGTCTTCTGCGTCC	8040
DB			
QY	8041	AACGAGAAAGGAGCGCGCAAGCTCGCTTATCGTATTCAGAGACTGGAGTTC	8100
DB			
QY	8041	AACGAGAAAGGAGCGCGCAAGCTCGCTTATCGTATTCAGAGACTGGAGTTC	8100
DB			
QY	8101	GTGTATCGGAGAGATGGCGCTTTACGACGTGTCTCCACCTTCTCAGCGCGTGTATGG	8160
DB			
QY	8101	GTGTATCGGAGAGATGGCGCTTTACGACGTGTCTCCACCTTCTCAGCGCGTGTATGG	8160
DB			
QY	8161	GCTCCTCATACGATTTCAATCTCCCCAAAGCAGCGGGTCGAGTTCCTGGTGATACCT	8220
DB			
QY	8161	GCTCCTCATACGATTTCAATCTCCCCAAAGCAGCGGGTCGAGTTCCTGGTGATACCT	8220
DB			
QY	8221	GGAAATCAAGAAATGCGCTTATGGGCTTCTCATATGACACCCGCTGTTTGACTCAACGG	8280
DB			
QY	8221	GGAAATCAAGAAATGCGCTTATGGGCTTCTCATATGACACCCGCTGTTTGACTCAACGG	8280
DB			
QY	8281	TCACTGAGTGACATTCGTTGAGGAGTCAATTTTACCAATGTTGTGACTTGGCCCCCG	8340
DB			
QY	8281	TCACTGAGTGACATTCGTTGAGGAGTCAATTTTACCAATGTTGTGACTTGGCCCCCG	8340
DB			
QY	8341	AGGCCAGACAGGCCATTAAGGTGCTCAGAGCGGCTTTTACATCGGGGGTCCCTGACTA	8400

8341 AGGCCAGACAGCCATAAGTTCGCTCACAGAGCGCTTTATCATCGGGGTCCCTCGACTA 8400
8401 ACTCAAAAGGCGAGAACTGCGGTATATCGCGGTGCGCGCAAGTGGCGTGTGACGACTA 8460
8401 ACTCAAAAGGCGAGAACTGCGGTATATCGCGGTGCGCGCAAGTGGCGTGTGACGACTA 8460
8461 GCTGCGGTAAATACCCCTCACATGTTACTTAAAGGCCACTGCGAGCCTGTGCGAGCTGCAAAAGC 8520
8461 GCTGCGGTAAATACCCCTCACATGTTACTTAAAGGCCACTGCGAGCCTGTGCGAGCTGCAAAAGC 8520
8521 TCCAGGACTGACAGATGCTGCGTGAACGAGAGCACTTGTGCTTATCTGTGAAAGCGCGG 8580
8521 TCCAGGACTGACAGATGCTGCGTGAACGAGAGCACTTGTGCTTATCTGTGAAAGCGCGG 8580
8581 GAACCCAGGAGATGCGCGGCCCTTACGAGCCTTACCGAGGCTATGACTAGGTATTCCG 8640
8581 GAACCCAGGAGATGCGCGGCCCTTACGAGCCTTACCGAGGCTATGACTAGGTATTCCG 8640
8641 CCCCCCGGGGATCCGCCCAACAGAAATACGACTGAGCTGATTAACATCATGTTCC 8700
8641 CCCCCCGGGGATCCGCCCAACAGAAATACGACTGAGCTGATTAACATCATGTTCC 8700
8701 CCAATGTGTGAGTCCGACAGATGATCTGGCAAAAGGATATACCTCACCCGTGACC 8760
8701 CCAATGTGTGAGTCCGACAGATGATCTGGCAAAAGGATATACCTCACCCGTGACC 8760
8761 CCACCAACCCCTTGCACGGGCTGCTGGGAGACAGCTAGACACACTCAATCAACTCTT 8820
8761 CCACCAACCCCTTGCACGGGCTGCTGGGAGACAGCTAGACACACTCAATCAACTCTT 8820
8821 GCTAGGCAATATCATGATGATGCGCCACCTATGGCAAGGATGATCTGTGATGACTC 8880
8821 GCTAGGCAATATCATGATGATGCGCCACCTATGGCAAGGATGATCTGTGATGACTC 8880
8881 ACTTTTCTCCATCTCTTAGCTCAAGAGCACTTGAAGGCGCTGATGTCAGATCT 8940
8881 ACTTTTCTCCATCTCTTAGCTCAAGAGCACTTGAAGGCGCTGATGTCAGATCT 8940
8941 ACGGGGCTGTACTCCATGAGGCACTTGACTACCTCAGATCAATGAAGCTCCATG 9000
8941 ACGGGGCTGTACTCCATGAGGCACTTGACTACCTCAGATCAATGAAGCTCCATG 9000
9001 GTCTTAGCGCATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT 9060
9001 GTCTTAGCGCATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT 9060
9061 GCCTCAGGAACTTGGGGTACCACCTTTCGCAACCTGGAGACATCGGGCCAGAGTGTC 9120
9061 GCCTCAGGAACTTGGGGTACCACCTTTCGCAACCTGGAGACATCGGGCCAGAGTGTC 9120
9121 GCGCTAAGCTACTGTCCAGGGGGGAGCGGCCACTTGTGGGAGATACCTCTTTAACT 9180
9121 GCGCTAAGCTACTGTCCAGGGGGGAGCGGCCACTTGTGGGAGATACCTCTTTAACT 9180
9181 GGGCAGTAAAGCAAGCTTAACTCACTCCAACTCCGCGCGCTCCAGCTGGACTTGT 9240
9181 GGGCAGTAAAGCAAGCTTAACTCACTCCAACTCCGCGCGCTCCAGCTGGACTTGT 9240
9241 CTGGCTGGTTCGTCGCTGTTTACAGCGGGGAGACATATATCACGCTGTCTCGTGCCC 9300
9241 CTGGCTGGTTCGTCGCTGTTTACAGCGGGGAGACATATATCACGCTGTCTCGTGCCC 9300
9301 GACCCGCTGTTTCCGTTGTCCTACTCTTCTGTTAGGGGTAGGCAATTTACCTGC 9360
9301 GACCCGCTGTTTCCGTTGTCCTACTCTTCTGTTAGGGGTAGGCAATTTACCTGC 9360
9361 TCCCAACCGATGAACGGGAGCTAACCACTCCAGGCTTAAAGCCATTTCTGTTTTTT 9420
9361 TCCCAACCGATGAACGGGAGCTAACCACTCCAGGCTTAAAGCCATTTCTGTTTTTT 9420
9421 TT 9480

9421 TT 9480
9481 TTTTTTCCCTTCTTTAATGGTGCTCCCATCTTAGCCCTAGTCAACGCTAGCTGTGAAAGCT 9540
9481 TTTTTTCCCTTCTTTAATGGTGCTCCCATCTTAGCCCTAGTCAACGCTAGCTGTGAAAGCT 9540
9541 CCGTGAGCCGCGCATGACTGCAGAGAGTCTGTATCTATCTGCGCTCTCTGCAGATCATGT 9595
9541 CCGTGAGCCGCGCATGACTGCAGAGAGTCTGTATCTATCTGCGCTCTCTGCAGATCATGT 9595
RESULT 2
AAC86939
ID AAC86939 standard; DNA; 9595 BP.
XX AAC86939;
AC AAC86939;
XX
DT 02-APR-2001 (first entry)
XX
DE Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 1b.
XX
KW Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
KW HCV; vaccine; viral inhibitor; antiviral; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..9374
FT /*tag= a
XX
FN WO200075352-A2.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000MO-US015527.
XX
PR 04-JUN-1999; 99US-0137817P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nam J, Bukh J, Emerson SU, Purcell RH;
XX
DR WPI; 2001-071081/08.
XX
P-PSDB; AAB31170.
XX
PT New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome
PT in which the (non-)structural region has been replaced by hepatitis C
PT virus (HCV) genome useful for treating or preventing HCV signs and
PT symptoms.
XX
PS Disclosure; Fig 4A-F; 97pp; English.
XX
CC The specification describes a nucleic acid comprising a chimeric virus
CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
CC the (non-)structural region has been replaced by the (non-)structural
CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC the chimeric virus and the chimeric virus are useful for identifying cell
CC lines capable of supporting the replication of these chimeric viruses, in
CC screening for neutralizing antibodies to HCV of different genotypes, in
CC the production of HCV-BVDV virions, for the development of inactivated or
CC attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
CC molecular properties of HCV indirectly in vitro, and in identifying
CC inhibitors of viral enzyme activity which would be useful as antiviral
CC agents. Formulations or compositions comprising the chimeric viruses may
CC be used to treat or prevent the signs and symptoms of HCV. The present
CC sequence represents a HCV clone, which is used to construct chimeric
CC nucleic acids of the invention
XX
SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;
Query Match 100.0%; Score 9595; DB 4; Length 9595;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2161 |||||GCCTAGTAGACTACCATACAGGCTTTGGCACTACCCCTGCACCTCAATTTTCCATCT2220
Qy 2221 TTAAGTTTAGATGATGTGGGGCGTGGAGCAAGAGCTCAATGCGCATGCAATTGGA2280
Db 2221 TTAAGTTTAGATGATGTGGGGCGTGGAGCAAGAGCTCAATGCGCATGCAATTGGA2280
Qy 2281 CTCGAGGAGCGCTGTAACCTTGAGGACAGGATAGGTCAAGACTCAGCCCGCTGCTGC2340
Db 2281 CTCGAGGAGAGCGCTGTAACTTTGGAGGACAGGATAGGTCAAGACTCAGCCCGCTGCTGC2340
Qy 2341 TGTCTACAAAGAGTGGCAGATACCTGCCCTGTGCTTTTCAACACCCCTACCGGCTTTATCCA2400
Db 2341 TGTCTACAAAGAGTGGCAGATACCTGCCCTGTGCTTTTCAACACCCCTACCGGCTTTATCCA2400
Qy 2401 CTGGTTTGATCTCATCTCCATCAGAAACATCGTGGAGTGCATACCTGTACGGGTAGGGT2460
Db 2401 CTGGTTTGATCCATCTCCATCAGAAACATCGTGGAGTGCATACCTGTACGGGTAGGGT2460
Qy 2461 CAGCGTTGTCTCCTTTGCAATCAATGGAGTACATCCTGTGCTTTTCTCTCTCTCTCG2520
Db 2461 CAGCGTTGTCTCCTTTGCAATCAATGGAGTACATCCTGTGCTTTTCTCTCTCTCTCG2520
Qy 2521 CAGACCGCGGTGTGCTGCTTTGTGATGATGCTGCTGATAGCCAGGCTGAGGCGG2580
Db 2521 CAGACCGCGGTGTGCTGCTTTGTGATGATGCTGCTGATAGCCAGGCTGAGGCGG2580
Qy 2581 CTTTAGAACTTGTGTGCTCTCAATGCGGCTCCGTGGCGGAGCGCATGTATTCCT2640
Db 2581 CTTTAGAACTTGTGTGCTCTCAATGCGGCTCCGTGGCGGAGCGCATGTATTCCT2640
Qy 2641 CTTTCTTGTGTTCTTCTGCGCGCTGGTACATTAAGGCGAGGCTGCTCCTGGGCGG2700
Db 2641 CTTTCTTGTGTTCTTCTGCGCGCTGGTACATTAAGGCGAGGCTGCTCCTGGGCGG2700
Qy 2701 CGTATGCTTTTATGCGGTATGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT2760
Db 2701 CGTATGCTTTTATGCGGTATGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT2760
Qy 2761 CTTACGCCCTTGACCGGAGATGGTGCATCGTGGGGGTGCGGTTCTTGTAGTCTGG2820
Db 2761 CTTACGCCCTTGACCGGAGATGGTGCATCGTGGGGGTGCGGTTCTTGTAGTCTGG2820
Qy 2821 TATTCTTGACCTGTCAACATFACTAAAGTGTCTCTACTAGGCTCATATGGTGGTTAC2880
Db 2821 TATTCTTGACCTGTCAACATFACTAAAGTGTCTCTACTAGGCTCATATGGTGGTTAC2880
Qy 2881 AATACTTTTATCACCAGAGCGGAGCGCATGCAAGTGTGGGTCCCGCCCTCAACGTTTC2940
Db 2881 AATACTTTTATCACCAGAGCGGAGCGCATGCAAGTGTGGGTCCCGCCCTCAACGTTTC2940
Qy 2941 GGGAGGCGCGATGCCATCATCTCTCTCACTGTGCGGTTTCATCCAGAGTTAAATTTTG3000
Db 2941 GGGAGGCGCGATGCCATCATCTCTCTCACTGTGCGGTTTCATCCAGAGTTAAATTTTG3000
Qy 3001 ACATACCAAACTCCTGCTGCGCATACTCGGCCGCTCATGTGCTCCAGGCTGSCATAA3060
Db 3001 ACATACCAAACTCCTGCTGCGCATACTCGGCCGCTCATGTGCTCCAGGCTGSCATAA3060
Qy 3061 CGAGAGTCCGCTACTTCTGCGCGCTCAAGGCTCATTCGTGCATGATGTTAGTGCAGTA3120
Db 3061 CGAGAGTCCGCTACTTCTGCGCGCTCAAGGCTCATTCGTGCATGATGTTAGTGCAGTA3120
Qy 3121 AAGTCCCGGGGTCAATTATGTCCAAATGGTCTTTCATGAAGCTGGCGGCTGACAGGTA3180
Db 3121 AAGTCCCGGGGTCAATTATGTCCAAATGGTCTTTCATGAAGCTGGCGGCTGACAGGTA3180
Qy 3181 CGTAGCTTTTAAACATCTTACCCATCGGGGACTGGGCCCAAGCGGCGCTACAGACC3240
Db 3181 CGTAGCTTTTAAACATCTTACCCATCGGGGACTGGGCCCAAGCGGCGCTACAGACC3240
Qy 3241 TTTGCGGTGGCGTAGAGCCGCTGCTCTCTCGCCATGGAGCCAAAGTTCATCCTCTGG3300
Db |||||

Db 3241 TTTGCGGTGGCGTAGAGCCCGTCTGTTCTTCCGCATGGAGCAAGGTCAATCACTCGG3300
Qy 3301 GAGCAGACACCGCTGCTGTGGGACATCATCTTGGGTCTACCGTCTCCGCCGGAAGG3360
Db 3301 GAGCAGACACCGCTGCTGTGGGACATCATCTTGGGTCTACCGTCTCCGCCGGAAGG3360
Qy 3361 GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGCAAGGGTGGGACTCTCTTG3420
Db 3361 GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGCAAGGGTGGGACTCTCTTG3420
Qy 3421 CGCCCATCAGGCCCTACTCCCAACAAACGCGGGCGTACTTTGGTTGATCATCATAGCC3480
Db 3421 CGCCCATCAGGCCCTACTCCCAACAAACGCGGGCGTACTTTGGTTGATCATCATAGCC3480
Qy 3481 TCACAGCCCGGCAAGAACAGGTCCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC3540
Db 3481 TCACAGCCCGGCAAGAACAGGTCCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC3540
Qy 3541 AATCTTTCTGGCGACCTGCATCAACGCGGTGCTGCTGACTGTCTTACCATGGCGTGGCT3600
Db 3541 AATCTTTCTGGCGACCTGCATCAACGCGGTGCTGCTGACTGTCTTACCATGGCGTGGCT3600
Qy 3601 CGAAGACCTTAGCCGCTCCAAAAGGTCCAATCAACCAATGTACCAATGTAGACCTGG3660
Db 3601 CGAAGACCTTAGCCGCTCCAAAAGGTCCAATCAACCAATGTACCAATGTAGACCTGG3660
Qy 3661 ACCTCGTGGGTGGCAGGCGCCCGGGGGCGCTCCATGACACCACTGAGCTGGCA3720
Db 3661 ACCTCGTGGGTGGCAGGCGCCCGGGGGCGCTCCATGACACCACTGAGCTGGCA3720
Qy 3721 GCTCGGACCTTTACTTGGTACGAGACATGTGATGTCAATTCGGTGGCGCGGAGGCG3780
Db 3721 GCTCGGACCTTTACTTGGTACGAGACATGTGATGTCAATTCGGTGGCGCGGAGGCG3780
Qy 3781 ACAGCAGGGGAAGTCTACTCTCCCGCCAGGCCGCTCTCTACTGAAAGGCTCTCGGGTG3840
Db 3781 ACAGCAGGGGAAGTCTACTCTCCCGCCAGGCCGCTCTCTACTGAAAGGCTCTCGGGTG3840
Qy 3841 GTCATTTGCTTTGCTCTCGGGGACGCTGCTGGGGTCTTCCGGCTGCTGTGTCACCC3900
Db 3841 GTCATTTGCTTTGCTCTCGGGGACGCTGCTGGGGGCTCTTCCGGCTGCTGTGTCACCC3900
Qy 3901 GGGGGTCCGAAGCGGTGACTTCATACCCGTTGAGTCTATGGAAACTACCATGCGGT3960
Db 3901 GGGGGTCCGAAGCGGTGACTTCATACCCGTTGAGTCTATGGAAACTACCATGCGGT3960
Qy 3961 CTCGGGTCTTCACAGACAACTCAACCCCGCGGTGTACCGCAGACATTTCCAAGTGGCAC4020
Db 3961 CTCGGGTCTTCACAGACAACTCAACCCCGCGGTGTACCGCAGACATTTCCAAGTGGCAC4020
Qy 4021 ATCTGCAGCTCTTACTGGCAGCGGACAGCACCAAGTCCGGCTGCGTATGCGAGCC4080
Db 4021 ATCTGCAGCTCTTACTGGCAGCGGACAGCACCAAGTCCGGCTGCGTATGCGAGCC4080
Qy 4081 AAGGTACAAGTGTCTGCTGTAACCCGTCCTGTAACCCGTCGTTGCGCACCTTAGGGTTGGGCGT4140
Db 4081 AAGGTACAAGTGTCTGCTGTAACCCGTCCTGTAACCCGTCGTTGCGCACCTTAGGGTTGGGCGT4140
Qy 4141 ATATGTCCAAGGCACACGGTATCGACCTTAACATCAGAACTGGGGTAAGGACCATTAACA4200
Db 4141 ATATGTCCAAGGCACACGGTATCGACCTTAACATCAGAACTGGGGTAAGGACCATTAACA4200
Qy 4201 CGGGCGCTCAATTACGTACTCCACCTATGGCAAGTTCCTTGGCCAGCGGTGGTCTGTCG4260
Db 4201 CGGGCGCTCAATTACGTACTCCACCTATGGCAAGTTCCTTGGCCAGCGGTGGTCTGTCG4260
Qy 4261 GGGGCGCTCATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACTACCATCT4320
Db 4261 GGGGCGCTCATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACTACCATCT4320
Qy 4321 TGGGCATCGGCACAGTCTCTGGACCAAGCGGAGACCGGTGGAGCGGCTCTGCTGCTCG4380
Db 4321 TGGGCATCGGCACAGTCTCTGGACCAAGCGGAGACCGGTGGAGCGGCTCTGCTGCTCG4380

QY	4381	CCACCGCTACACCTCCGGGATCGGTTACCGTGCACACCCCAATATCGAGGAAATAGGCC	4440
DB	4381	CCACCGCTACACCTCCGGGATCGGTTACCGTGCACACCCCAATATCGAGGAAATAGGCC	4440
QY	4441	TGTCGAACAAATGGAGATCCCTCTTATGGCAAGCCATCCCAATGAGGCCATCAAGG	4500
DB	4441	TGTCGAACAAATGGAGATCCCTCTTATGGCAAGCCATCCCAATGAGGCCATCAAGG	4500
QY	4501	GGGGAGGCACTCTCATTTCTGCAATCCAAAGAAATGTGACGAGCTCGCGCAAAAGC	4560
DB	4501	GGGGAGGCACTCTCATTTCTGCCATCCAAAGAAATGTGACGAGCTCGCGCAAAAGC	4560
QY	4561	TGACAGCCTCGAAGCTGAGCTGAGCATATTAACCGGGCCCTTGTATGTCTCGTATAC	4620
DB	4561	TGACAGCCTCGAAGCTGAGCTGAGCATATTAACCGGGCCCTTGTATGTCTCGTATAC	4620
QY	4621	CGCCTATCGGAGCGTGTGTGCTGCGCAACAGACGCTCTAATGACGGGTTTCACCGGG	4680
DB	4621	CGCCTATCGGAGCGTGTGTGCTGCGCAACAGACGCTCTAATGACGGGTTTCACCGGG	4680
QY	4681	ATTTTGACTCAGTGATCGACTGCAATACATGTGTCAACAGACAGTCGACTTCAGCTTGG	4740
DB	4681	ATTTTGACTCAGTGATCGACTGCAATACATGTGTCAACAGACAGTCGACTTCAGCTTGG	4740
QY	4741	ATCCCACTTACCAATTTGAGACGACGACCGTGCACCAAGACGCGGTGTGCGCTCGGAAC	4800
DB	4741	ATCCCACTTACCAATTTGAGACGACGACCGTGCACCAAGACGCGGTGTGCGCTCGGAAC	4800
QY	4801	GGCGAGTAGAATGCGCAGGGGTAGAGTGGCATCTACAGTTTGTGACTCCAGGAGAAC	4860
DB	4801	GGCGAGTAGAATGCGCAGGGGTAGAGTGGCATCTACAGTTTGTGACTCCAGGAGAAC	4860
QY	4861	GGCCCTCGGCGATGTTCTGATTTCTGCTGCTGTGTGAGTCTATGACGCGGCTGTGCTT	4920
DB	4861	GGCCCTCGGCGATGTTCTGATTTCTGCTGCTGTGTGAGTCTATGACGCGGCTGTGCTT	4920
QY	4921	GGTATGAGCTCAGCCCGCTGAGACCTCGTGGTGGTGGGCTTACCTAAATACACCAG	4980
DB	4921	GGTATGAGCTCAGCCCGCTGAGACCTCGTGGTGGTGGGCTTACCTAAATACACCAG	4980
QY	4981	GGTTGCCCTGTGCGAGGACCATCTGAGTCTTGGGAGAGCGTCTTACAGGCTCACCC	5040
DB	4981	GGTTGCCCTGTGCGAGGACCATCTGAGTCTTGGGAGAGCGTCTTACAGGCTCACCC	5040
QY	5041	ACATAGATGCCACTTCTGCTCCAGACTAAACAGGAGAGACAACTTTCTTACCTGG	5100
DB	5041	ACATAGATGCCACTTCTGCTCCAGACTAAACAGGAGAGACAACTTTCTTACCTGG	5100
QY	5101	TGGCATATCAAGCTTACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCAATCGTGGACCAAA	5160
DB	5101	TGGCATATCAAGCTTACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCAATCGTGGACCAAA	5160
QY	5161	TGTGGAAGTGTCTATACGGCTGAAACTTACATGTGACGGGCAACCCCTGCTGTATA	5220
DB	5161	TGTGGAAGTGTCTATACGGCTGAAACTTACATGTGACGGGCAACCCCTGCTGTATA	5220
QY	5221	GGCTAGGAGCGCTCCAAATGAGGTCTCTTCAACACCCCATAACTAAATACATATGG	5280
DB	5221	GGCTAGGAGCGCTCCAAATGAGGTCTCTTCAACACCCCATAACTAAATACATATGG	5280
QY	5281	CATGCATGTCCGCTGACCTGAGGTCTGTCACTAGCACCTGGGTCTGGTAGGCGAGTCC	5340
DB	5281	CATGCATGTCCGCTGACCTGAGGTCTGTCACTAGCACCTGGGTCTGGTAGGCGAGTCC	5340
QY	5341	TTGAGCTTTGGCGCATACTCTGTGACGACAGGAGTGGGTCAATGTGGGAGGATCA	5400
DB	5341	TTGAGCTTTGGCGCATACTCTGTGACGACAGGAGTGGGTCAATGTGGGAGGATCA	5400
QY	5401	TCCTTGTCCGGAAGCGCTGTGCTTCCGACAGGAGTCTCTACAGGAGTTCGATG	5460
DB	5401	TCCTTGTCCGGAAGCGCTGTGCTTCCGACAGGAGTCTCTACAGGAGTTCGATG	5460

QY	5461	AGATCGAAGAGTGTGCTTCAAACTTCTTATCATCGAGCAGGAAATGACAGTCTCGCGAGC	5520
DB	5461	AGATCGAAGAGTGTGCTTCAAACTTCTTATCATCGAGCAGGAAATGACAGTCTCGCGAGC	5520
QY	5521	AATTTCAAGCAAAAGCGCTCGGGTTGTGCAAAACGGCCACCAAGCAAGCGAGGCTGCTG	5580
DB	5521	AATTTCAAGCAAAAGCGCTCGGGTTGTGCAAAACGGCCACCAAGCAAGCGAGGCTGCTG	5580
QY	5581	CTCCGCTGTGAGTCCAAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCACATGGA	5640
DB	5581	CTCCGCTGTGAGTCCAAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCACATGGA	5640
QY	5641	ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAACCCCGCGA	5700
DB	5641	ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAACCCCGCGA	5700
QY	5701	TAGCATCATTTGATGGCATTTACAGCTTTTATCATAGCCGCTCACCAACCAAAACACC	5760
DB	5701	TAGCATCATTTGATGGCATTTACAGCTTTTATCATAGCCGCTCACCAACCAAAACACC	5760
QY	5761	TCCTGTTTAAACATCTTTGGGGGATGGTGGCTGCCCCAACTCGCTCTCCAGCGCTGCGT	5820
DB	5761	TCCTGTTTAAACATCTTTGGGGGATGGTGGCTGCCCCAACTCGCTCTCCAGCGCTGCGT	5820
QY	5821	CAGCTTTCTGCGGCGCGCATCGCGGAGCGCTGTGGCAGCATAGGCTTTGGGAAGG	5880
DB	5821	CAGCTTTCTGCGGCGCGCATCGCGGAGCGCTGTGGCAGCATAGGCTTTGGGAAGG	5880
QY	5881	TGCTCGTGACATCTTTGGGGGCTATGGGGCAGGGGTAGCGGCGCATCTCGTGGCCTTTA	5940
DB	5881	TGCTCGTGACATCTTTGGGGGCTATGGGGCAGGGGTAGCGGCGCATCTCGTGGCCTTTA	5940
QY	5941	AGTCTAGCGCGGAGGTGCCCTCCACGAGACCTGTGTCAACTTACTTCCCTGCCATCC	6000
DB	5941	AGTCTAGCGCGGAGGTGCCCTCCACGAGACCTGTGTCAACTTACTTCCCTGCCATCC	6000
QY	6001	TCCTCTCGTGCCCTGCTCGGGGTCTGTGCGCAGCAATACTGCGTGGCAGCTGG	6060
DB	6001	TCCTCTCGTGCCCTGCTCGGGGTCTGTGCGCAGCAATACTGCGTGGCAGCTGG	6060
QY	6061	GCCCGGAGAGGGGCTGTGCACTGATGAACCGGCTGATAGCGTTGCGTTCGGGGGTA	6120
DB	6061	GCCCGGAGAGGGGCTGTGCACTGATGAACCGGCTGATAGCGTTGCGTTCGGGGGTA	6120
QY	6121	ACCACGCTCTCCCTACGCACTATGTGCTGAGAGCGAGCGTGCAGCACTGTCTCAGA	6180
DB	6121	ACCACGCTCTCCCTACGCACTATGTGCTGAGAGCGAGCGTGCAGCACTGTCTCAGA	6180
QY	6181	TCCTCTAGCCTTACCATCACTCACTGCTGAAGCGGCTCCACCACTGGATTAATCAGG	6240
DB	6181	TCCTCTAGCCTTACCATCACTCACTGCTGAAGCGGCTCCACCACTGGATTAATCAGG	6240
QY	6241	ACTGCTCTAGCCATGCTCCGGCTCGTGGCTTAAAGGATGTTTGGGATTTGGATATGACGG	6300
DB	6241	ACTGCTCTAGCCATGCTCCGGCTCGTGGCTTAAAGGATGTTTGGGATTTGGATATGACGG	6300
QY	6301	TGTTGACTGACTTCAAGACCTGCTCCAGTCCAAACTCTCTGCGCGGTTACCGGGAGTCC	6360
DB	6301	TGTTGACTGACTTCAAGACCTGCTCCAGTCCAAACTCTCTGCGCGGTTACCGGGAGTCC	6360
QY	6361	CTTTCTGTGTCATGCCAACGCGGGTACAAAGGAGTCTGCGCGGGGAGCGGCATCATGAAA	6420
DB	6361	CTTTCTGTGTCATGCCAACGCGGGTACAAAGGAGTCTGCGCGGGGAGCGGCATCATGAAA	6420
QY	6421	CCACTGCCATCGGAGCACAGATCCCGGACATGTCAAAAAACGTTTCCATGAGGATCG	6480
DB	6421	CCACTGCCATCGGAGCACAGATCCCGGACATGTCAAAAAACGTTTCCATGAGGATCG	6480
QY	6481	TAGGCGCTAGAACCTGCGACCAACAGTGGGACCGGAAACGTTTCCCATCAACGATACCA	6540
DB	6481	TAGGCGCTAGAACCTGCGACCAACAGTGGGACCGGAAACGTTTCCCATCAACGATACCA	6540
QY	6541	CGGAGACCTTGACACACCTCCCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGGTGGCTG	6600

Db 6541 |||||CGGGACCTTGCACACCTCCCGGGGCCCAACTATTCCAGGGCGCTATGGCGGTGGCTG 6600
Qy 6601 CTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGCATGACCA 6660
Db 6601 CTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGCATGACCA 6660
Qy 6661 CTGACAAAGTAAAGTGCCTCCATGCCAGGTTCCGGCCCCCGAAATCTTTACGGAGGTGGATG 6720
Db 6661 CTGACAAAGTAAAGTGCCTCCATGCCAGGTTCCGGCCCCCGAAATCTTTACGGAGGTGGATG 6720
Qy 6721 GAGTGGGTTGCACAGGTACGCTCCGGCGTGCAAACTCTTCTACGGAGGACGTCAAGT 6780
Db 6721 GAGTGGGTTGCACAGGTACGCTCCGGCGTGCAAACTCTTCTACGGAGGACGTCAAGT 6780
Qy 6781 TCCAGGTGGGCTCAACCAATACTTGGTGGGTCCAGCTCCCATCGAGCCCGAACCGG 6840
Db 6781 TCCAGGTGGGCTCAACCAATACTTGGTGGGTCCAGCTCCCATCGAGCCCGAACCGG 6840
Qy 6841 AGCTAACAGTCTTACTTCCATGCTCACCGATCCCTCCACATTAACAGCAGAGACGGCTA 6900
Db 6841 AGCTAACAGTCTTACTTCCATGCTCACCGATCCCTCCACATTAACAGCAGAGACGGCTA 6900
Qy 6901 AGCGTAGGCTGCTAGAGGTTCTCCCTCTTTTAGCCAGCTCATCAGTACGCCAGTTGT 6960
Db 6901 AGCGTAGGCTGCTAGAGGTTCTCCCTCTTTTAGCCAGCTCATCAGTACGCCAGTTGT 6960
Qy 6961 CTGCGCTCTTTGAAGGCGATCGCATPACCCACCATGACTCCCGGACGCTGACCTCA 7020
Db 6961 CTGCGCTCTTTGAAGGCGATCGCATPACCCACCATGACTCCCGGACGCTGACCTCA 7020
Qy 7021 TCGAGGCCAACTCTTGTGGCGGAGAGATGGGGGAAACATCACTCGCGTGGAGTCAG 7080
Db 7021 TCGAGGCCAACTCTTGTGGCGGAGAGATGGGGGAAACATCACTCGCGTGGAGTCAG 7080
Qy 7081 AGAAATAAGGTAGTAATCTTGGAATCTTTGCAACCGCTTCAACGGAGGGGATGAGAGG 7140
Db 7081 AGAAATAAGGTAGTAATCTTGGAATCTTTGCAACCGCTTCAACGGAGGGGATGAGAGG 7140
Qy 7141 AGATATCCGTCCGGCGGAGATCTCGGNAATCCAGNAAGTTCCTCAGGGTTGCCCA 7200
Db 7141 AGATATCCGTCCGGCGGAGATCTCGGNAATCCAGNAAGTTCCTCAGGGTTGCCCA 7200
Qy 7201 TATGGCAGCCCGGACTACAACTCTCACTAGAGTCTGAGGACCGGAGACCGGACTAG 7260
Db 7201 TATGGCAGCCCGGACTACAACTCTCACTAGAGTCTGAGGACCGGAGACCGGACTAG 7260
Qy 7261 TCCCTCCGGTGATACAGGATGCCATTTGCCACCTACCAAGGCTCCTCCAAATACCACCTC 7320
Db 7261 TCCCTCCGGTGATACAGGATGCCATTTGCCACCTACCAAGGCTCCTCCAAATACCACCTC 7320
Qy 7321 CACGAGAAAGGACGGTTGTCTGACAGAAATCCAAATGCTGCTTCTGCGTTGGCGGAGC 7380
Db 7321 CACGAGAAAGGACGGTTGTCTGACAGAAATCCAAATGCTGCTTCTGCGTTGGCGGAGC 7380
Qy 7381 TCGCCACTAAGACCTTCGGTAGCTCCGGATCGTGGCGGTGATAGCGSCACGGCGACCG 7440
Db 7381 TCGCCACTAAGACCTTCGGTAGCTCCGGATCGTGGCGGTGATAGCGSCACGGCGACCG 7440
Qy 7441 CCCTTCTGACCTTCCGACGAGCGGTGACAAAGGATCCGACGTTGAGTCGTACTCCT 7500
Db 7441 CCCTTCTGACCTTCCGACGAGCGGTGACAAAGGATCCGACGTTGAGTCGTACTCCT 7500
Qy 7501 CCATGCCCTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGGCTTTGGTCTA 7560
Db 7501 CCATGCCCTTGAAGGGAGCGGGGAGCCCGGATCTCAGCGACGGGCTTTGGTCTA 7560
Qy 7561 CCGTGAGTGAGAGGCTAGTAGGATGTCGTCTGCTGCTCAATGCTTATAGCTGACAG 7620
Db 7561 CCGTGAGTGAGAGGCTAGTAGGATGTCGTCTGCTGCTCAATGCTTATAGCTGACAG 7620
Qy 7621 GCGCCTGTATCGGCATCGCTGCGGAGGAAGTAAAGTGCCTCAACCCGTTGAGCA 7680
Db 7621 GCGCCTGTATCGGCATCGCTGCGGAGGAAGTAAAGTGCCTCAACCCGTTGAGC

Db 7621 GCGCCTGTATCGGCATCGCTGCGGAGGAAGTAAAGTGCCTCAACCCGTTGAGCA 7680
Qy 7681 ACTCTTTGCTCGTCAACCAACATGTGTCTACGCCACAACATCCCGCAGCGCAAGCTCC 7740
Db 7681 ACTCTTTGCTCGTCAACCAACATGTGTCTACGCCACAACATCCCGCAGCGCAAGCTCC 7740
Qy 7741 GGCAGAAAGGTCACTTTTCACAGATTTGCAAGTCTCTGGATGATCATTTACCGGACGTAC 7800
Db 7741 GGCAGAAAGGTCACTTTTCACAGATTTGCAAGTCTCTGGATGATCATTTACCGGACGTAC 7800
Qy 7801 TCAAGGAGATGAAGCGAGCGGTCCACAGTTAAGGCTAAGCTTCTATCTATAGAGGAG 7860
Db 7801 TCAAGGAGATGAAGCGAGCGGTCCACAGTTAAGGCTAAGCTTCTATCTATAGAGGAG 7860
Qy 7861 CCTGCAAGCTCAGCCCCCACAATTCGGCCAAATTTGGCTATGGGGCAAGGACG 7920
Db 7861 CCTGCAAGCTCAGCCCCCACAATTCGGCCAAATTTGGCTATGGGGCAAGGACG 7920
Qy 7921 TCCGAACTTATCCAGCAGGGCCGTTAACCATCTCCGCTCCGTTGGGAGGACTTGTGG 7980
Db 7921 TCCGAACTTATCCAGCAGGGCCGTTAACCATCTCCGCTCCGTTGGGAGGACTTGTGG 7980
Qy 7981 AAGACATGAAACCAATTTGACACCAACATCATGGCAAAAGTGAAGTTTCTCGCTCC 8040
Db 7981 AAGACATGAAACCAATTTGACACCAACATCATGGCAAAAGTGAAGTTTCTCGCTCC 8040
Qy 8041 AACCAGAGAAGGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGACCTGGGAGTTC 8100
Db 8041 AACCAGAGAAGGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGACCTGGGAGTTC 8100
Qy 8101 GTGTATGCGAGAAGATGSCCTTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGG 8160
Db 8101 GTGTATGCGAGAAGATGSCCTTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGG 8160
Qy 8161 GCTCTCTATACGATTTCAATTAATCTCCCAAGCAGCGGTGAGTTCTGTGTGAATACCT 8220
Db 8161 GCTCTCTATACGATTTCAATTAATCTCCCAAGCAGCGGTGAGTTCTGTGTGAATACCT 8220
Qy 8221 GGAAATCAAAAGATGCCCCCTATGGCTTCTCATATGACACCGCGTGTGTGACTCAACGG 8280
Db 8221 GGAAATCAAAAGATGCCCCCTATGGCTTCTCATATGACACCGCGTGTGTGACTCAACGG 8280
Qy 8281 TCACCTGAGAGTGACATTCGTGTAGGAGTCAATTTACCAATGTGTGACTTGGCCCCCG 8340
Db 8281 TCACCTGAGAGTGACATTCGTGTAGGAGTCAATTTACCAATGTGTGACTTGGCCCCCG 8340
Qy 8341 AGGCAGACAGCCATAAGTCTGCTCACAGACGGCTTTATATCGGGGTCCCTGACTA 8400
Db 8341 AGGCAGACAGCCATAAGTCTGCTCACAGACGGCTTTATATCGGGGTCCCTGACTA 8400
Qy 8401 ACTCAAAAGGGCAGAACTGCGGTTATCGCGGTCCCGCAGTGGCGTGTGACGACTA 8460
Db 8401 ACTCAAAAGGGCAGAACTGCGGTTATCGCGGTCCCGCAGTGGCGTGTGACGACTA 8460
Qy 8461 GCTCGGTAAATACCTCACAATTTACTTTGAAGGCCACTTCAGCCCTGTGAGCTGCAAAAGC 8520
Db 8461 GCTCGGTAAATACCTCACAATTTACTTTGAAGGCCACTTCAGCCCTGTGAGCTGCAAAAGC 8520
Qy 8521 TCCAGGACTGACGATGCTGCTGAAACGGAGACGCTTGTCTGTTATCTGTGAAAGCGCGG 8580
Db 8521 TCCAGGACTGACGATGCTGCTGAAACGGAGACGCTTGTCTGTTATCTGTGAAAGCGCGG 8580
Qy 8581 GAAACCCAGGAGATGCGGCGGCTTACGAGCTTACGAGGCTTACGAGGCTATGACTAGTATTCG 8640
Db 8581 GAAACCCAGGAGATGCGGCGGCTTACGAGCTTACGAGGCTTACGAGGCTATGACTAGTATTCG 8640
Qy 8641 CCCCCCGGGGATCCGCCCCCAACCAAGATACGACTCTGGAGCTGATAACATCATGTTCCT 8700
Db 8641 CCCCCCGGGGATCCGCCCCCAACCAAGATACGACTCTGGAGCTGATAACATCATGTTCCT 8700
Qy 8701 CCAATGTGTAGTCCGCGACGATCTGGGCAAGGATGATCTGAGTAAAGGTTATACCTCAACCGGTGACC 8760
Db 8701 CCAATGTGTAGTCCGCGACGATCTGGGCAAGGTTATACCTCAACCGGTGACC 8760

```
QY 8761 CCACCAACCCCTTGCACGGGCTGCGTGGGAGACAGCTAGACACACTCCAAATCAACTCTT 8820
DB 8761 CCACCAACCCCTTGCACGGGCTGCGTGGGAGACAGCTAGACACACTCCAAATCAACTCTT 8820
QY 8821 GGTAGGCATATCATATGATATGCGCCACCTATGGCAAGATGATCTGATGACTC 8880
DB 8821 GGTAGGCATATCATATGATATGCGCCACCTATGGCAAGATGATCTGATGACTC 8880
QY 8881 ACTTTTCTCCATCCCTTAGCTCAAGAGCAACTTGAAAAGCCCTGGATTGTGAGATCT 8940
DB 8881 ACTTTTCTCCATCCCTTAGCTCAAGAGCAACTTGAAAAGCCCTGGATTGTGAGATCT 8940
QY 8941 ACGGGCTTCTACTCATTTAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCATG 9000
DB 8941 ACGGGCTTCTACTCATTTAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCATG 9000
QY 9001 GTCTTAGCGATTTACACTCCAGTTTACTCTCAGGTGAGATCAATAGGGTGGCTTCAT 9060
DB 9001 GTCTTAGCGATTTACACTCCAGTTTACTCTCAGGTGAGATCAATAGGGTGGCTTCAT 9060
QY 9061 GCCTCAGGAACCTGGGGTACCACTTGGAACTCGGACATCGGGCCAGAAAGTGTC 9120
DB 9061 GCCTCAGGAACCTGGGGTACCACTTGGAACTCGGACATCGGGCCAGAAAGTGTC 9120
QY 9121 GCGCTAAGCTACTGTCCAGGGGGAGGGCCGCACTTGTGGCAGATACCTCTTTAACT 9180
DB 9121 GCGCTAAGCTACTGTCCAGGGGGAGGGCCGCACTTGTGGCAGATACCTCTTTAACT 9180
QY 9181 GGGCAGTAAGGACCAAGCTTTAACTACTCAATCCCGGGCGGTCCAGCTGAGATTGT 9240
DB 9181 GGGCAGTAAGGACCAAGCTTTAACTACTCAATCCCGGGCGGTCCAGCTGAGATTGT 9240
QY 9241 CTGGCTGGTTCGTGCTGTACAGCGGGGAGACATATATCAGACCTCTCTCGTGCC 9300
DB 9241 CTGGCTGGTTCGTGCTGTACAGCGGGGAGACATATATCAGACCTCTCTCGTGCC 9300
QY 9301 GACCCGCTGTTTCCGTTGCTTACTCTTACTTCTGTAGGGGTAGGCAATTTACCTGC 9360
DB 9301 GACCCGCTGTTTCCGTTGCTTACTCTTACTTCTGTAGGGGTAGGCAATTTACCTGC 9360
QY 9361 TCCCAACCGATGAACGGGAGCTAACCACTCCAGGCTTTAAGCAATTTCTGTTTTTT 9420
DB 9361 TCCCAACCGATGAACGGGAGCTAACCACTCCAGGCTTTAAGCAATTTCTGTTTTTT 9420
QY 9421 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9480
DB 9421 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9480
QY 9481 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9540
DB 9481 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9540
QY 9541 CCGTGAGCGGATGACCTGCAGAGAGTGTGATAGTGGCTCTCTGCGAGATCATGT 9595
DB 9541 CCGTGAGCGGATGACCTGCAGAGAGTGTGATAGTGGCTCTCTGCGAGATCATGT 9595

RESULT 3
ID ADO36222 standard; DNA; 9595 BP.
XX ADO36222;
AC AC
XX AC
XX AC
DT 26-AUG-2004 (first entry)
XX Hepatitis C virus (HCV) J4L6 wild-type cDNA.
DE hepatotropic; virucide; vaccine; gene therapy; vaccine;
XX Hepatitis C virus; HCV; core protein; HCV infection; vaccination; gene;
KW ds.
XX Hepatitis C virus.
OS
```

```
XX Key Location/Qualifiers
FH CDS 342..9374
FT /*tag= a
FT /product= "HCV polypeptide"
XX WO2004046175-A1.
XX 03-JUN-2004.
XX 13-NOV-2003; 2003WO-EP012793.
XX 15-NOV-2002; 2002GB-00026722.
XX (GLAXO) GLAXO GROUP LTD.
XX Brett S, Hamblin PA, Ogilvie L;
XX WPI; 2004-420613/39.
XX P-PSDB; ADO36227.
XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT the polypeptide sequences of the HCV core and at least one other HCV
PT protein, for use in medicine, particularly for manufacturing a medicament
PT for treating HCV.
XX Disclosure; Fig 1; 78pp; English.
XX The invention describes a polynucleotide vaccine comprising a
CC polynucleotide sequence (S1) encoding the Hepatitis C virus (HCV) Core
CC protein and at least 1 other HCV protein, and causes expression of the
CC proteins in cells in which (S1) has been mutated or positioned relative
CC to the polynucleotide sequence encoding the other HCV protein, so that
CC the negative effect of the Core protein on expression of the other HCV
CC protein is reduced). Also described are: a method of preventing or
CC treating an HCV infection in a mammal, comprising administering of an
CC vaccine cited above to a mammal; and a method of vaccination of an
CC individual, comprising taking a polynucleotide vaccine as cited above,
CC coating the polynucleotide onto gold beads and delivering the gold beads
CC into the skin. HCV nucleic acids, polypeptides, host cells, vectors and
CC antibodies used in the methods, are also disclosed. The polynucleotide
CC vaccine is useful in the manufacture of a medicament for the treatment of
CC HCV. This sequence encodes the wild type HCV polypeptide.
XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;
```

```
Query Match 100.0%; Score 9595; DB 12; Length 9595;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCCCCCTGATGGGGGCGACACTCCACATGAATCACTCCCTCTGTAGGAATCACTG 60
DB 1 GCCAGCCCCCTGATGGGGGCGACACTCCACATGAATCACTCCCTCTGTAGGAATCACTG 60
QY 61 TCTTACGCAAGAAAGCGTCTAGCCATGGCGTGTAGTATGATGATGCTGAGCCTCCAGGAC 120
DB 61 TCTTACGCAAGAAAGCGTCTAGCCATGGCGTGTAGTATGATGATGCTGAGCCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTGCCAG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTGCCAG 180
QY 181 GACGACCGGGTCTTTCTTTGGATCAACCCGCTCAATGCTGGAGATTTGGCGTGGCCCC 240
DB 181 GACGACCGGGTCTTTCTTTGGATCAACCCGCTCAATGCTGGAGATTTGGCGTGGCCCC 240
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCGCTTGTGTGTACTGCCCTGATAGG 300
DB 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCGCTTGTGTGTACTGCCCTGATAGG 300
QY 301 GTGCTTCGAGTGGCCCGGGAGGTCTGTAGACCGGTGCACCATGACGACCAATCTTAAC 360
DB 301 GTGCTTCGAGTGGCCCGGGAGGTCTGTAGACCGGTGCACCATGACGACCAATCTTAAC 360
```

QY 361 CTCAAAGAAAACCAAAAGTAAACAAACCGCGGCCCAAGAGCTCAAGTTCCCGGCG 420
DB 361 CTCAAAGAAAACCAAAAGTAAACAAACCGCGGCCCAAGAGCTCAAGTTCCCGGCG 420
QY 421 GTGGTCAGATCGTTGGTGGAGTTTACCTGTGTCCCGCAGAGGGCCCCAGGTTGGGTGTC 480
DB 421 GTGGTCAGATCGTTGGTGGAGTTTACCTGTGTCCCGCAGAGGGCCCCAGGTTGGGTGTC 480
QY 481 GCGGACTAGGAAGGCTTCCGAGCGGTGCGAACCTCGTGGGAAGGGCACAACCTATCCCAA 540
DB 481 GCGGACTAGGAAGGCTTCCGAGCGGTGCGAACCTCGTGGGAAGGGCACAACCTATCCCAA 540
QY 541 AGGCTCGCCGACCCGAGGGCAGGGCTGGGCTCAGCCGCGGTACCGCTTGGCCCTCTATG 600
DB 541 AGGCTCGCCGACCCGAGGGCAGGGCTGGGCTCAGCCGCGGTACCGCTTGGCCCTCTATG 600
QY 601 GCAATGAGGGCTGGGGTGGGAGGATGGCTCTGTCAACCCGCGCTCCCGGCTAGTT 660
DB 601 GCAATGAGGGCTGGGGTGGGAGGATGGCTCTGTCAACCCGCGCTCCCGGCTAGTT 660
QY 661 GGGGCCCCAGGACCCCGCGGTAGGTGCGGTAACTTGGGTAAGTCAATCGATACCCCTTA 720
DB 661 GGGGCCCCAGGACCCCGCGGTAGGTGCGGTAACTTGGGTAAGTCAATCGATACCCCTTA 720
QY 721 CATGGGCTTCGCGATCTCATGGGGTACATTCGCTCGTCGCGCGCCCTAGGGGGCG 780
DB 721 CATGGGCTTCGCGATCTCATGGGGTACATTCGCTCGTCGCGCGCCCTAGGGGGCG 780
QY 781 CTGCCAGGGCTTGGCACACGGTGTCCGGGTTCTGGAGGACCGGTGAACATGCAACAG 840
DB 781 CTGCCAGGGCTTGGCACACGGTGTCCGGGTTCTGGAGGACCGGTGAACATGCAACAG 840
QY 841 GGAACTTGCCCGGTGCTCTTCTCTATCTCTTCTTCTTGGCTCTGTCTCTTCTTGGACCA 900
DB 841 GGAACTTGCCCGGTGCTCTTCTCTATCTCTTCTTCTTGGCTCTGTCTCTTCTTGGACCA 900
QY 901 TCCGAGCTTCGCTTATGAAGTGGCGAACGCTCGCGGATATACCATGTCACGACGACT 960
DB 901 TCCGAGCTTCGCTTATGAAGTGGCGAACGCTCGCGGATATACCATGTCACGACGACT 960
QY 961 GCTCCAACTCAAGCATTGTATGAGGAGCGGAAGTATCATGCAATCTCCCGGGTGG 1020
DB 961 GCTCCAACTCAAGCATTGTATGAGGAGCGGAAGTATCATGCAATCTCCCGGGTGG 1020
QY 1021 TGCCTGTGTTGAGGAGGTAAACAGCTCCGTTGTGGGTAGGCTCACTCCACGCTCG 1080
DB 1021 TGCCTGTGTTGAGGAGGTAAACAGCTCCGTTGTGGGTAGGCTCACTCCACGCTCG 1080
QY 1081 CCGCCAGGAATGCCAGGTCCTCACTACGACATAGCAGCCACGTCGACTTGTCTGCTTG 1140
DB 1081 CCGCCAGGAATGCCAGGTCCTCACTACGACATAGCAGCCACGTCGACTTGTCTGCTTG 1140
QY 1141 GGAAGGCTGCTTCTGCTCCGCTATGATAGTGGGGATCTCTGCGGATCTATTTTCTCG 1200
DB 1141 GGAAGGCTGCTTCTGCTCCGCTATGATAGTGGGGATCTCTGCGGATCTATTTTCTCG 1200
QY 1201 TCTCCAGCTGTTCACCTTCTGCTCCGCTGCGGATAGAGACAGTGCAGGACTGCAACTGCT 1260
DB 1201 TCTCCAGCTGTTCACCTTCTGCTCCGCTGCGGATAGAGACAGTGCAGGACTGCAACTGCT 1260
QY 1261 CAATCTATCCCGGCAATGATCAGGTACCGCATGGCTTGGGATATGATGTAAGTGGT 1320
DB 1261 CAATCTATCCCGGCAATGATCAGGTACCGCATGGCTTGGGATATGATGTAAGTGGT 1320
QY 1321 CAGCTACACAGCCCTAGTGGTGTGCGAGTTGCTCCGGATCCACAAAGCTGTCGTGGACA 1380
DB 1321 CAGCTACACAGCCCTAGTGGTGTGCGAGTTGCTCCGGATCCACAAAGCTGTCGTGGACA 1380
QY 1381 TGTGGCGGGGCCCACTGGGAGTCTTGGCGGCTTGGCTACTATTCATGGTAGGGA 1440
DB 1381 TGTGGCGGGGCCCACTGGGAGTCTTGGCGGCTTGGCTACTATTCATGGTAGGGA 1440

QY 1441 ACTGGGCTAAGGTTCTGATTGTGGCGCTACTCTTTGCGGGCGTTGACGGGAGAGCCACA 1500
DB 1441 ACTGGGCTAAGGTTCTGATTGTGGCGCTACTCTTTGCGGGCGTTGACGGGAGAGCCACA 1500
QY 1501 CGACGGGAGGGTGGCGGCCACACCACTCCGGGTTACGTCCTCTTTCTCATCTGGGG 1560
DB 1501 CGACGGGAGGGTGGCGGCCACACCACTCCGGGTTACGTCCTCTTTCTCATCTGGGG 1560
QY 1561 CGTCTCAGAAAATCAGCTTGTGAATACCAACGCGAGCTGCACATCAACAGGACTGCC 1620
DB 1561 CGTCTCAGAAAATCAGCTTGTGAATACCAACGCGAGCTGCACATCAACAGGACTGCC 1620
QY 1621 TAAATTCGAATGACTCCCTCCAAACTGGGTTCTTTGCGCGCTGTTTACGCACAAGT 1680
DB 1621 TAAATTCGAATGACTCCCTCCAAACTGGGTTCTTTGCGCGCTGTTTACGCACAAGT 1680
QY 1681 TCAACTGTCGGGTGCCCGAGGCGATGGCAGCTCGCGGCCCATTTGACTGGTTGGGCC 1740
DB 1681 TCAACTGTCGGGTGCCCGAGGCGATGGCAGCTCGCGGCCCATTTGACTGGTTGGGCC 1740
QY 1741 AGGGTGGGGCCCCATCACCTATATTAAGCCTAACAGCTCGSATCAGAGGCTTATTGCT 1800
DB 1741 AGGGTGGGGCCCCATCACCTATATTAAGCCTAACAGCTCGSATCAGAGGCTTATTGCT 1800
QY 1801 GGCATTACGCGCTCGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
DB 1801 GGCATTACGCGCTCGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
QY 1861 ATTGTTTCACCCCAAGCCCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
DB 1861 ATTGTTTCACCCCAAGCCCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
QY 1921 ATAGCTGGGGGAGNAATGAGACAGCTGATGCTCTCAACAAACAGCGTCCGCCACAAG 1980
DB 1921 ATAGCTGGGGGAGNAATGAGACAGCTGATGCTCTCTCAACAAACAGCGTCCGCCACAAG 1980
QY 1981 GCAACTGTTCCGGCTGTACATGATGAATAGTACTTGGGTTCACTAAGACGTGCGGAGTCT 2040
DB 1981 GCAACTGTTCCGGCTGTACATGATGAATAGTACTTGGGTTCACTAAGACGTGCGGAGTCT 2040
QY 2041 CCCCCTGTAAACATCGGGGGGTGCGTAAACCGCACCTTGATCTGCCCCACGGAAGTCTTCC 2100
DB 2041 CCCCCTGTAAACATCGGGGGGTGCGTAAACCGCACCTTGATCTGCCCCACGGAAGTCTTCC 2100
QY 2101 GGAAGCACCCGAGGCTACTTACAAAATGTGGTCTCGGGCCCTGTTGTGACACCTAGT 2160
DB 2101 GGAAGCACCCGAGGCTACTTACAAAATGTGGTCTCGGGCCCTGTTGTGACACCTAGT 2160
QY 2161 GCCTAGTAGACTTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTCCATCT 2220
DB 2161 GCCTAGTAGACTTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTCCATCT 2220
QY 2221 TTAAGGTTAGATGATGTGGGGGCGTGGAGCACAGGCTCAATGCGCATGCAATGGA 2280
DB 2221 TTAAGGTTAGATGATGTGGGGGCGTGGAGCACAGGCTCAATGCGCATGCAATGGA 2280
QY 2281 CTCGAGGAGGCGCTGTAACTTGGAGGACAGGATAGTCAAGACTCAGCCCGCTGCTCG 2340
DB 2281 CTCGAGGAGGCGCTGTAACTTGGAGGACAGGATAGTCAAGACTCAGCCCGCTGCTCG 2340
QY 2341 TGTCTACAAACAGAGTGGCAGATACTGCCCCTGTGCTTTACCAACCTACCGGCTTATCCA 2400
DB 2341 TGTCTACAAACAGAGTGGCAGATACTGCCCCTGTGCTTTACCAACCTACCGGCTTATCCA 2400
QY 2401 CTGGTTTGCATCCATCTCCATCAGAAACATCGTGGACGTCGTAATCTGTACGCTAGGCT 2460
DB 2401 CTGGTTTGCATCCATCTCCATCAGAAACATCGTGGACGTCGTAATCTGTACGCTAGGCT 2460
QY 2461 CAGCGCTTGTCTCTTTCGAATCAATGGGAGTATCTCTGTTGTTTCTTCTTCTCTCG 2520
DB 2461 CAGCGCTTGTCTCTTTCGAATCAATGGGAGTATCTCTGTTGTTTCTTCTTCTCTCG 2520
QY 2521 CAGACGCGCGGCTGTGTGCTTGTGGATGATGCTGTGTATAGCCAGGCTGAGGCCG 2580

Db 6901 ||||| AGCGTAGGCTGGCTAGAGGCTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTTGT 6960
Qy 6961 CTGCGGCTTTCTTTGAAGGCGACATGCATACCCACCATGACTCCCCGGAGCGTGACTCA 7020
Db 6961 CTGCGGCTTTCTTTGAAGGCGACATGCATACCCACCATGACTCCCCGGAGCGTGACTCA 7020
Qy 7021 TCAGAGCCAACTCTTGTGGCGGACAGAGATGGCGGAAACATCATCTCGGTGGAGTCAG 7080
Db 7021 TCAGAGCCAACTCTTGTGGCGGACAGAGATGGCGGAAACATCATCTCGGTGGAGTCAG 7080
Qy 7081 AGAATAAGGTAGTAATCTCGACTCTTTGAAACCGCTTCACGCGGAGGGGATGAGAGG 7140
Db 7081 AGAATAAGGTAGTAATCTCGACTCTTTGAAACCGCTTCACGCGGAGGGGATGAGAGG 7140
Qy 7141 AGATATCCGTGGCGGAGATCCTCGGAAATCCAGGAAGTTCCCTCAGCGTTGCCCA 7200
Db 7141 AGATATCCGTGGCGGAGATCCTCGGAAATCCAGGAAGTTCCCTCAGCGTTGCCCA 7200
Qy 7201 TATGGGACGCCCGGACTCAATCTCTCACTGTAGAGTCTCGAAGGACCCGAGCTACG 7260
Db 7201 TATGGGACGCCCGGACTCAATCTCTCACTGTAGAGTCTCGAAGGACCCGAGCTACG 7260
Qy 7261 TCCCTCGGTGTACACGGATGCCATTCGACCTACCAAGGCTCCTCCAATACCACTC 7320
Db 7261 TCCCTCGGTGTACACGGATGCCATTCGACCTACCAAGGCTCCTCCAATACCACTC 7320
Qy 7321 CACGGAAGAGAGACGGTTGTCTGACAGAATCCAATGTCTTCTGCTTGGCGGAGC 7380
Db 7321 CACGGAAGAGAGACGGTTGTCTGACAGAATCCAATGTCTTCTGCTTGGCGGAGC 7380
Qy 7381 TCGCCATAAGACTTCGGTAGCTCCGGATCGTGCGCGCTTGATAGCGGACGCGACCG 7440
Db 7381 TCGCCATAAGACTTCGGTAGCTCCGGATCGTGCGCGCTTGATAGCGGACGCGACCG 7440
Qy 7441 CCCTTCTGACTGGCTCCGACGACGGTGACAAAGATCCGAGCTTGTAGTCTGACTCCT 7500
Db 7441 CCCTTCTGACTGGCTCCGACGACGGTGACAAAGATCCGAGCTTGTAGTCTGACTCCT 7500
Qy 7501 CCATGCCCCCTTGAAGGGAGCGGGGACCCGATCTCAGGACGGTCTTGGTCTA 7560
Db 7501 CCATGCCCCCTTGAAGGGAGCGGGGACCCGATCTCAGGACGGTCTTGGTCTA 7560
Qy 7561 CCGTAGTGAGGAGCTAGTGAGGATGTCTGCTCTCAATGTCTCTATACGTGGACAG 7620
Db 7561 CCGTAGTGAGGAGCTAGTGAGGATGTCTGCTCTCAATGTCTCTATACGTGGACAG 7620
Qy 7621 GCGCCCTGATCAGCCATGCGCTCGGAGGAAGTAAGCTGCCCATCAACCCGTTGAGCA 7680
Db 7621 GCGCCCTGATCAGCCATGCGCTCGGAGGAAGTAAGCTGCCCATCAACCCGTTGAGCA 7680
Qy 7681 ACTCTTTGCTGCTCACCACAATGTCTACGCCCAACATCCCGACGCGACCTCC 7740
Db 7681 ACTCTTTGCTGCTCACCACAATGTCTACGCCCAACATCCCGACGCGACCTCC 7740
Qy 7741 GGCAAGAAAGCTCACCTTTGACAGATTGCAAGTCTCGGATGATCAATTACCGGACGTAC 7800
Db 7741 GGCAAGAAAGCTCACCTTTGACAGATTGCAAGTCTCGGATGATCAATTACCGGACGTAC 7800
Qy 7801 TCAAGGAGATGAAGCGGAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGG 7860
Db 7801 TCAAGGAGATGAAGCGGAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGG 7860
Qy 7861 CCTGCAAGCTGACGCCCCCAATTTCGGCAATCCAAATTTGGCTATGGGCAAGGACG 7920
Db 7861 CCTGCAAGCTGACGCCCCCAATTTCGGCAATCCAAATTTGGCTATGGGCAAGGACG 7920
Qy 7921 TCCGGAACCTATCAGAGGCGGTAAACACATCCGCTCGCTGCGGAGACTTGTGG 7980
Db 7921 TCCGGAACCTATCAGAGGCGGTAAACACATCCGCTCGCTGCGGAGACTTGTGG 7980
Qy 7981 AAGACACTGAACCACTTACACCACTCATGCGCAAAAGTGAGTTTCTCGCTCC 8040

Db 7981 AAGACACTGAACCACTTACACCACTCATGCGCAAAAGTGAGTTTCTCGCTCC 8040
Qy 8041 AACGAGAGAGGAGCGCGCAAGCCAGCTCGCTTATCGTATTCAGACCTGGGAGTTC 8100
Db 8041 AACGAGAGAGGAGCGCGCAAGCCAGCTCGCTTATCGTATTCAGACCTGGGAGTTC 8100
Qy 8101 GTGTATGCGAGAAAGATGGCCCTTTAGCAGTGTCTCCACCTTCTCAGGCGCGTATGG 8160
Db 8101 GTGTATGCGAGAAAGATGGCCCTTTAGCAGTGTCTCCACCTTCTCAGGCGCGTATGG 8160
Qy 8161 GCTCTCTCATACGGAATTCATAACTCTCCCAAGCAGCGGTTCGAGTTCTTGGTGAATACCT 8220
Db 8161 GCTCTCTCATACGGAATTCATAACTCTCCCAAGCAGCGGTTCGAGTTCTTGGTGAATACCT 8220
Qy 8221 GGAAATCAAAAGAAATGCGCTTATGAGAGTCAATTTACCAATGTTGTGACTTGGCCCCCG 8280
Db 8221 GGAAATCAAAAGAAATGCGCTTATGAGAGTCAATTTACCAATGTTGTGACTTGGCCCCCG 8280
Qy 8281 TCACTGAGAGTGACATTCGTTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCG 8340
Db 8281 TCACTGAGAGTGACATTCGTTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCG 8340
Qy 8341 AGGCCAGACAGGCCATAAGGTCTCAGAGCGGCTTTACATCGGGGTCCCCTGACTA 8400
Db 8341 AGGCCAGACAGGCCATAAGGTCTCAGAGCGGCTTTACATCGGGGTCCCCTGACTA 8400
Qy 8401 ACTCAAAAGGCGAAGACTCGGTTATCGCGGTTCGCGCGCAAGTGGCGTCTGACGACTA 8460
Db 8401 ACTCAAAAGGCGAAGACTCGGTTATCGCGGTTCGCGCGCAAGTGGCGTCTGACGACTA 8460
Qy 8461 GCTGCGGTAAATACCTTACATGTTACTGAAAGCCACTGCGAGCTCTGCGAGCTGCAAGC 8520
Db 8461 GCTGCGGTAAATACCTTACATGTTACTGAAAGCCACTGCGAGCTCTGCGAGCTGCAAGC 8520
Qy 8521 TCCAGGACTGACAGTCTCGTGAAAGCGAGACGACCTTGTCTTACTGTGAAAGCGCGG 8580
Db 8521 TCCAGGACTGACAGTCTCGTGAAAGCGAGACGACCTTGTCTTACTGTGAAAGCGCGG 8580
Qy 8581 GAAACCCAGGAGGATGCGCGGCGCTTACGAGCTTCAAGGCTTCAAGGCTTATGACTTATTCG 8640
Db 8581 GAAACCCAGGAGGATGCGCGGCGCTTACGAGCTTCAAGGCTTCAAGGCTTATGACTTATTCG 8640
Qy 8641 CCCCCCGGGGATCGCGCCCAACAGAAATACGACCTGGAGCTGATAACATCATGTTCCT 8700
Db 8641 CCCCCCGGGGATCGCGCCCAACAGAAATACGACCTGGAGCTGATAACATCATGTTCCT 8700
Qy 8701 CCAATGTGTGAGTCTGCGGCGCATGCGTCTGCAAAAGGGTATACTACTCACCCGTCACC 8760
Db 8701 CCAATGTGTGAGTCTGCGGCGCATGCGTCTGCAAAAGGGTATACTACTCACCCGTCACC 8760
Qy 8761 CCAACCCCGCTTTCGACCGGCTCGGTGGGAGACAGCTAGACACACTCCAATCAACTCTT 8820
Db 8761 CCAACCCCGCTTTCGACCGGCTCGGTGGGAGACAGCTAGACACACTCCAATCAACTCTT 8820
Qy 8821 GGCTAGGCAATATCATGTATGCGGCCCAACCTATGGGCAAGGATGATTCGTAGTACTC 8880
Db 8821 GGCTAGGCAATATCATGTATGCGGCCCAACCTATGGGCAAGGATGATTCGTAGTACTC 8880
Qy 8881 ACTTTTCTCCATCTTCTAGCTCAAGGCAACTTGAAGGCGCTCGATTTGTGAGATCT 8940
Db 8881 ACTTTTCTCCATCTTCTAGCTCAAGGCAACTTGAAGGCGCTCGATTTGTGAGATCT 8940
Qy 8941 ACGGGGTCTGCTACTCTTCAATGAGCCACTTACCTACTCTCAGATCATTTGAACGACTCCATG 9000
Db 8941 ACGGGGTCTGCTACTCTTCAATGAGCCACTTACCTACTCTCAGATCATTTGAACGACTCCATG 9000
Qy 9001 GTCTTAGCGCATTTACACTCCAAGTTACTCTCAGGTGAGATCAATAGGGTGGCTTCAT 9060
Db 9001 GTCTTAGCGCATTTACACTCCAAGTTACTCTCAGGTGAGATCAATAGGGTGGCTTCAT 9060
Qy 9061 GCCTCAGAAACTTGGGGTACCCTTGGGACCTTGGAGCATCGGGCCAGAGTGTCC 9120
Db 9061 GCCTCAGAAACTTGGGGTACCCTTGGGACCTTGGAGCATCGGGCCAGAGTGTCC 9120


```
QY 9121 GCGCTAAGCTACTGTCCAGGGGGGAGGGCGGCACCTTGTGGCAGATACCTCTTTAACT 9180
DB 9121 GCGCTAAGCTACTGTCCAGGGGGGAGGGCGGCACCTTGTGGCAGATACCTCTTTAACT 9180
QY 9181 GGGCAGTAAGGACCAAGCTTAAACTCACTCCAAATCCCGGCGCGTCCAGCTGGACTTGT 9240
DB 9181 GGGCAGTAAGGACCAAGCTTAAACTCACTCCAAATCCCGGCGCGTCCAGCTGGACTTGT 9240
QY 9241 CTGGCTGTTCTGCTCGTGTACAGGGGGGAGACATATATACAGCCTGTCTGTGCCC 9300
DB 9241 CTGGCTGTTCTGCTCGTGTACAGGGGGGAGACATATATACAGCCTGTCTGTGCCC 9300
QY 9301 GACCCCGCTGGTTCCTGTTGTGCTACTCTCTACTTCTGTAGGGGTAGGCATTTACCTGC 9360
DB 9301 GACCCCGCTGGTTCCTGTTGTGCTACTCTCTACTTCTGTAGGGGTAGGCATTTACCTGC 9360
QY 9361 TCCCAACCGATGAACGGGGAGCTAACCACTCCAGGCTTAAGCCATTTCCTGTTTTTTT 9420
DB 9361 TCCCAACCGATGAACGGGGAGCTAACCACTCCAGGCTTAAGCCATTTCCTGTTTTTTT 9420
QY 9421 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9480
DB 9421 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9480
QY 9481 TTTTTTCCCTCTTTAATGTGGCTCCATCTTAGCCCTAGTCAAGGCTAGCTGTGAAGGT 9540
DB 9481 TTTTTTCCCTCTTTAATGTGGCTCCATCTTAGCCCTAGTCAAGGCTAGCTGTGAAGGT 9540
QY 9541 CCGTGAGCGCATGACTGCAGAGAGTGTGTGACTTGGCTCTCTCTGAGATCATGT 9595
DB 9541 CCGTGAGCGCATGACTGCAGAGAGTGTGTGACTTGGCTCTCTCTGAGATCATGT 9595

RESULT 4
ID ADO79396
XX ADO79396 standard; cDNA; 9595 BP.
AC ADO79396;
XX
XX
XX 26-AUG-2004 (first entry)
DE Hepatitis C virus J4L6 genome wild-type cDNA.
DE
KW HCV; polyprotein; vaccine; DNA immunisation; hepatotropic; virucide;
KW gene; ss.
XX
XX Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 342..9374
XX FT /*tag= a
XX FT /product= "HCV polyprotein"
XX
XX WO2004046176-A1.
XX
XX 03-JUN-2004.
XX
XX 13-NOV-2003; 2003WO-EP012830.
XX
XX 15-NOV-2002; 2002GB-00026722.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Brett S, Hamblin PA, Ogilvie L;
XX
XX WPI; 2004-420614/39.
DR P-PSDS; ADO79401.
DR GENBANK; AF054247.
XX
XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,
```

for use in medicine, in particular for manufacturing a medicament for the treatment of HCV.

PS Disclosure; Fig 1; 79pp; English.

XX The present sequence is the hepatitis C virus (HCV) J4L6 genome wild-type
CC cDNA sequence encoding the HCV polyprotein. HCV DNA vaccines of the
CC invention comprise a polynucleotide that encodes the HCV proteins Core,
CC NS3, NS4B and NS5B, and does not encode the NS4A and/or NS5A proteins.
CC Preferably, the codon usage of the polynucleotide sequence resembles that
CC of highly expressed human genes. The polynucleotides may encode
CC individual proteins or fusion proteins. Preferred fusions include double
CC fusions between NS4B and NS5B and between Core and NS3. The vaccines are
CC useful for the treatment or prevention of an HCV infection.

XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;

Query Match 100.0%; Score 9595; DB 12; Length 9595;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GCCAGCCCCCTGATGGGGGCGACACTCCACATCAATCACTCCCTCTGTGAGAACTACTG 60
DB 1 GCCAGCCCCCTGATGGGGGCGACACTCCACATCAATCACTCCCTCTGTGAGAACTACTG 60
QY 61 TCCTTCAGCGAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTGAGCCTCCAGGAC 120
DB 61 TCCTTCAGCGAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTGAGCCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATTCGCG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATTCGCG 180
QY 181 GACGACGGGTCTTTCTTGATCAACCCGCTCAATGCTGGAGATTGGCGTGCCTCC 240
DB 181 GACGACGGGTCTTTCTTGATCAACCCGCTCAATGCTGGAGATTGGCGTGCCTCC 240
QY 241 CGGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGGCCCTTGTGCTACTGCTGATAGG 300
DB 241 CGGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGGCCCTTGTGCTACTGCTGATAGG 300
QY 301 GTGCTTGGAGTGTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAAGAACTCTAAAC 360
DB 301 GTGCTTGGAGTGTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAAGAACTCTAAAC 360
QY 361 CTCAGAGAAACCAAGCTTAACCAACCCGCGCCACAGAGCGTCAAGTTCCTCCGGCG 420
DB 361 CTCAGAGAAACCAAGCTTAACCAACCCGCGCCACAGAGCGTCAAGTTCCTCCGGCG 420
QY 421 GTGGTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCAGGGGCCAGAGTTGGGTGTGC 480
DB 421 GTGGTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCAGGGGCCAGAGTTGGGTGTGC 480
QY 481 CGCGACTTAGGAAGCTTCCGAGCGGTCCAACTCTGTGGAAGGCGACAACATATCCCAA 540
DB 481 CGCGACTTAGGAAGCTTCCGAGCGGTCCAACTCTGTGGAAGGCGACAACATATCCCAA 540
QY 541 AGGCTCGCCGACCCGAGGGCAGGGCTGGGTTCAGCCCGGGTACCTTTGGCCCTCTATG 600
DB 541 AGGCTCGCCGACCCGAGGGCAGGGCTGGGTTCAGCCCGGGTACCTTTGGCCCTCTATG 600
QY 601 GCAATGAGGGCTTGGGTGGGCGAGGATGGCTCTGTCTACCCCGGGCTCCCGGCTAGTT 660
DB 601 GCAATGAGGGCTTGGGTGGGCGAGGATGGCTCTGTCTACCCCGGGCTCCCGGCTAGTT 660
QY 661 GGGGCCCCACGGAACCCCGGCGTAGGTTCGCTAACTTGGGTAAAGTCAATCGATACCTTA 720
DB 661 GGGGCCCCACGGAACCCCGGCGTAGGTTCGCTAACTTGGGTAAAGTCAATCGATACCTTA 720
QY 721 CATGGGCTTGGCGGATCTCATGGGTACATTCGCTCGGCGCCCCCTAGGGGGCG 780
DB 721 CATGGGCTTGGCGGATCTCATGGGTACATTCGCTCGGCGCCCCCTAGGGGGCG 780
```

Qy	781	CTCCAGGGCTTGGACACGGTGTCCGGTTCTGGAGACGGCTGAATATGCAACAG	840
Db	781	CTGCCAGGGCTTGGACACGGTGTCCGGTTCTGGAGACGGCTGAATATGCAACAG	840
Qy	841	GGAACTTGGCGGTTGCTCTTTCTCTATCTTCTCTTGGCTCTGCTCTCTGTTGACCA	900
Db	841	GGAACTTGGCGGTTGCTCTTTCTCTATCTTCTCTTGGCTCTGCTCTCTGTTGACCA	900
Qy	901	TCCAGCTTCCGCTTATGAAGTGGCAACGCTGTCGGGATATACCAATGTCACGAAACGACT	960
Db	901	TCCAGCTTCCGCTTATGAAGTGGCAACGCTGTCGGGATATACCAATGTCACGAAACGACT	960
Qy	961	GCTCCAACTCAAGCATTTGTATGAGCGACGGACGTGATCATATGTCATPACTCCCGGTTGG	1020
Db	961	GCTCCAACTCAAGCATTTGTATGAGCGACGGACGTGATCATATGTCATPACTCCCGGTTGG	1020
Qy	1021	TGCCCTGTGTTACGAGGGTAAACAGTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCG	1080
Db	1021	TGCCCTGTGTTACGAGGGTAAACAGTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCG	1080
Qy	1081	CGGCCAGGAATGCCAGCTCCCACTACGACAAATACGACGCCAGCTGCTGCTGTTG	1140
Db	1081	CGGCCAGGAATGCCAGCTCCCACTACGACAAATACGACGCCAGCTGCTGCTGTTG	1140
Qy	1141	GGACGGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTCTGGGATCTATTTTCCTCG	1200
Db	1141	GGACGGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTCTGGGATCTATTTTCCTCG	1200
Qy	1201	TCTCCAGCTGTTACCTTCTCGCTCGCGCATGAGACAGTGCAGGACTGCAACTGCT	1260
Db	1201	TCTCCAGCTGTTACCTTCTCGCTCGCGCATGAGACAGTGCAGGACTGCAACTGCT	1260
Qy	1261	CAATCTATCCCGGCATGTATACAGTCAACGATGGCTTGGGATATGATGAATGACTGGT	1320
Db	1261	CAATCTATCCCGGCATGTATACAGTCAACGATGGCTTGGGATATGATGAATGACTGGT	1320
Qy	1321	CACCTACAAAGCCCTAGTGTGTCGAGTGTCTCCGATCCCAACAGCTGTCTGGACA	1380
Db	1321	CACCTACAAAGCCCTAGTGTGTCGAGTGTCTCCGATCCCAACAGCTGTCTGGACA	1380
Qy	1381	TGGTGGCGGGGCCACTGGGGAGTCTTGGCGGCTTGCCTACTATTTCCATGTAGGGA	1440
Db	1381	TGGTGGCGGGGCCACTGGGGAGTCTTGGCGGCTTGCCTACTATTTCCATGTAGGGA	1440
Qy	1441	ACTGGGCTAAGGTTCTGATTTGTGCGCTACTCTTTTGGCGGTTGACGGGAGACCCACA	1500
Db	1441	ACTGGGCTAAGGTTCTGATTTGTGCGCTACTCTTTTGGCGGTTGACGGGAGACCCACA	1500
Qy	1501	CGACGGGAGGGTGGCGGCCACACCACTCCGGGTTCACTCCCTTTCTCATCTGGGG	1560
Db	1501	CGACGGGAGGGTGGCGGCCACACCACTCCGGGTTCACTCCCTTTCTCATCTGGGG	1560
Qy	1561	CGTCTCAGAAATCCAGTTGTGAATACCAACGGCAGCTGGCACATCAACAGGACTGCC	1620
Db	1561	CGTCTCAGAAATCCAGTTGTGAATACCAACGGCAGCTGGCACATCAACAGGACTGCC	1620
Qy	1621	TAAATTCGAATGACTCCCTCCAACTGGGTTCTTTTGGCGGCTGTTTACGCACACAAGT	1680
Db	1621	TAAATTCGAATGACTCCCTCCAACTGGGTTCTTTTGGCGGCTGTTTACGCACACAAGT	1680
Qy	1681	TCAAATCTCCCGGGTGGCGGAGGCATGGCCAGCTGCGGCCCATTTGACTGGTTCGCC	1740
Db	1681	TCAAATCTCCCGGGTGGCGGAGGCATGGCCAGCTGCGGCCCATTTGACTGGTTCGCC	1740
Qy	1741	AGGGGTGGGGCCCCATCACTATATAAGCTAAACAGCTCGGATCAGAGGCTTATTGCT	1800
Db	1741	AGGGGTGGGGCCCCATCACTATATAAGCTAAACAGCTCGGATCAGAGGCTTATTGCT	1800
Qy	1801	GGCATTAAGCGCTCGACCGTGTGTTGCTACCGGCTCGCAGGTGTGGTCCAGTGT	1860
Db	1801	GGCATTAAGCGCTCGACCGTGTGTTGCTACCGGCTCGCAGGTGTGGTCCAGTGT	1860
Qy	1861	ATTGTTTACCCCAAGCCCTGTGTGGTGGGGACCAACCGATCGTTCGGTGTCCCTACGT	1920

Db	1861	ATTGTTTACCCCAAGCCCTGTGTGGTGGGGACCAACCGATCGTTCGGTGTCCCTACGT	1920
Qy	1921	ATAGCTGGGGGAGAAATGAGACAGAGTGTCTCTCAACACACAGCGCTCCGCCACAAG	1980
Db	1921	ATAGCTGGGGGAGAAATGAGACAGAGTGTCTCTCAACACACAGCGCTCCGCCACAAG	1980
Qy	1981	GCAACTGGTTCGGCTGTATCATGATGAATAGTACTGGGTTCACTAAGACGTGCGGAGTTC	2040
Db	1981	GCAACTGGTTCGGCTGTATCATGATGAATAGTACTGGGTTCACTAAGACGTGCGGAGTTC	2040
Qy	2041	CCCCGTGTAAACATCGGGGGGTCCGTAAACGGACCTTGATCTGCCACAGGACTGCTTC	2100
Db	2041	CCCCGTGTAAACATCGGGGGGTCCGTAAACGGACCTTGATCTGCCACAGGACTGCTTC	2100
Qy	2101	GGAAACACCCCGAGGCTACTTACACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT	2160
Db	2101	GGAAACACCCCGAGGCTACTTACACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT	2160
Qy	2161	GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTCCATCT	2220
Db	2161	GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTCCATCT	2220
Qy	2221	TTAAGGTTAGGATGTATGTGGGGCGTGGAGCACAGGCTCAATGCGGCATGCAATTTGGA	2280
Db	2221	TTAAGGTTAGGATGTATGTGGGGCGTGGAGCACAGGCTCAATGCGGCATGCAATTTGGA	2280
Qy	2281	CTCGAGGAGCGCTGTAACTTGGAGGACAGGATAGGTGAGAACTCAGCCGCTGCTGC	2340
Db	2281	CTCGAGGAGCGCTGTAACTTGGAGGACAGGATAGGTGAGAACTCAGCCGCTGCTGC	2340
Qy	2341	TGCTACACAGAGTGCAGATACTGCCCTGTCTTTTACACCCCTACCGGCTTTATCCA	2400
Db	2341	TGCTACACAGAGTGCAGATACTGCCCTGTCTTTTACACCCCTACCGGCTTTATCCA	2400
Qy	2401	CTGGTTTGATTCCTCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGTGTAGGGT	2460
Db	2401	CTGGTTTGATTCCTCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGTGTAGGGT	2460
Qy	2461	CAGCGTTTGTCTCTTTTGGCAATCAATGGAGTACATCTCTGTGTCTTCTCTCTGG	2520
Db	2461	CAGCGTTTGTCTCTTTTGGCAATCAATGGAGTACATCTCTGTGTCTTCTCTCTGG	2520
Qy	2521	CAGACGGCGCGTGTGCTGCTTGTGGATGATGCTGCTGATAGCCAGGCTGAGGCCG	2580
Db	2521	CAGACGGCGCGTGTGCTGCTTGTGGATGATGCTGCTGATAGCCAGGCTGAGGCCG	2580
Qy	2581	CCTTAGAGAACTTGGTGGTCTCAATGCGCGCTCCGTGGCGGAGCGCATGTTCTCT	2640
Db	2581	CCTTAGAGAACTTGGTGGTCTCAATGCGCGCTCCGTGGCGGAGCGCATGTTCTCT	2640
Qy	2641	CTTTCTTGTGTTCTTCTGCGCGCTGTGTACATTAAGGCGAGGCTGGCTCTCGGGCGG	2700
Db	2641	CTTTCTTGTGTTCTTCTGCGCGCTGTGTACATTAAGGCGAGGCTGGCTCTCGGGCGG	2700
Qy	2701	CGTATGCTTTTATGGCGTATGGCGCTGCTCTCTGCTCTTACTGCGGCTTACCAACGAG	2760
Db	2701	CGTATGCTTTTATGGCGTATGGCGCTGCTCTCTGCTCTTACTGCGGCTTACCAACGAG	2760
Qy	2761	CTTACGCTTTGGACCGGGAGATGGCTCATCTGTCGGGGGTGGGTTCTTGTAGTCTGG	2820
Db	2761	CTTACGCTTTGGACCGGGAGATGGCTCATCTGTCGGGGGTGGGTTCTTGTAGTCTGG	2820
Qy	2821	TATTTCTTGACTGTGTACCAAGTGTTCCTACTAGGCTCATATGGTGGTGTAC	2880
Db	2821	TATTTCTTGACTGTGTACCAAGTGTTCCTACTAGGCTCATATGGTGGTGTAC	2880
Qy	2881	AATACTTTTATCACCAGAGCGAGGCGCACATGCAAGTGTGGGTCCCCCCCCCTCAACGTTT	2940
Db	2881	AATACTTTTATCACCAGAGCGAGGCGCACATGCAAGTGTGGGTCCCCCCCCCTCAACGTTT	2940
Qy	2941	GGGAGGCGCGATGCCATCATCTCTCTCAGTGTGGGTTTCATCCAGAGTTAATTTTGT	3000

Dd 2941 GGGAGGCGCGGATGCCATCATCTCTCTCAAGTGTGCGGTTCCAGAGTAAATTTTG 3000
Qy 3001 ACATCAACAACTCTCTGCTCGCATACTCGGCCCTCATGTGTCTCCAGGCTGSCATAA 3060
Dd 3001 ACATCAACAACTCTCTGCTCGCATACTCGGCCCTCATGTGTCTCCAGGCTGSCATAA 3060
Qy 3061 CGAGAGTCCGCTGCTGTCGCGCTCAAGGCTCATTCGTGCATGCATGTTAGTGCAG 3120
Dd 3061 CGAGAGTCCGCTGCTGTCGCGCTCAAGGCTCATTCGTGCATGCATGTTAGTGCAG 3120
Qy 3121 AGTCCGCGGGGTCAATATGTCGAAATGCTTTCATGAAGTGGCGGCTGACAGGTA 3180
Dd 3121 AGTCCGCGGGGTCAATATGTCGAAATGCTTTCATGAAGTGGCGGCTGACAGGTA 3180
Qy 3181 CGTAGCTTTATAACCATCTTACCCACTCGGGACTGGGCCACGCGGCTTACGAGACC 3240
Dd 3181 CGTAGCTTTATAACCATCTTACCCACTCGGGACTGGGCCACGCGGCTTACGAGACC 3240
Qy 3241 TTGCGGTGGCGGTAGAGCCCGTCTTCTTCGCCCATGGAGCAAGGTTCATCCTGGG 3300
Dd 3241 TTGCGGTGGCGGTAGAGCCCGTCTTCTTCGCCCATGGAGCAAGGTTCATCCTGGG 3300
Qy 3301 GAGCAGACCCCTGGGTGGGACATCATCTTGGGTCTACCCGTCTCCGCCCGAAGG 3360
Dd 3301 GAGCAGACCCCTGGGTGGGACATCATCTTGGGTCTACCCGTCTCCGCCCGAAGG 3360
Qy 3361 GGAAGAGATATTTTGGGACCGGTGATGTCGAAAGGGCAAGGTGGGACTCCTTG 3420
Dd 3361 GGAAGAGATATTTTGGGACCGGTGATGTCGAAAGGGCAAGGTGGGACTCCTTG 3420
Qy 3421 CGCCCATCAGCGCTACTCCCAACAAACGCGGGGGTACTTGGTGGCATCATAGCC 3480
Dd 3421 CGCCCATCAGCGCTACTCCCAACAAACGCGGGGGTACTTGGTGGCATCATAGCC 3480
Qy 3481 TCACAGCCGGGACAGAAACAGGTGCAAGGGAGGTTCAAGTGTCTACCGCAACAC 3540
Dd 3481 TCACAGCCGGGACAGAAACAGGTGCAAGGGAGGTTCAAGTGTCTACCGCAACAC 3540
Qy 3541 AATCTTCTCGGACCTGTCATCAAGCGGTGTGGTGTCTTACCATGCGGTGGCT 3600
Dd 3541 AATCTTCTCGGACCTGTCATCAAGCGGTGTGGTGTCTTACCATGCGGTGGCT 3600
Qy 3601 CGAAGACCTTAGCGGTCCAAAGGTCCAAATCACCCAAATGTACCAATGTAGACCTGG 3660
Dd 3601 CGAAGACCTTAGCGGTCCAAAGGTCCAAATCACCCAAATGTAGACCTGG 3660
Qy 3661 ACTCGTGGGTGGAGGCGCCCGCGGGCGGTCCATGACACCATGCACTGTGGCA 3720
Dd 3661 ACTCGTGGGTGGAGGCGCCCGCGGGCGGTCCATGACACCATGCACTGTGGCA 3720
Qy 3721 GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCATTCGGGTGGCGGAGGG 3780
Dd 3721 GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCATTCGGGTGGCGGAGGG 3780
Qy 3781 ACAGAGGGGAAGTCTACTCTCCCGAGCGCTCTCTACCTCGAAAGGCTCCTCGGGTG 3840
Dd 3781 ACAGAGGGGAAGTCTACTCTCCCGAGCGCTCTCTACCTCGAAAGGCTCCTCGGGTG 3840
Qy 3841 GTCCATTTGCTTTGCCCTTCGGGGCAGCTCGTGGGGCTCTTCGGGGTGTGTGACCC 3900
Dd 3841 GTCCATTTGCTTTGCCCTTCGGGGCAGCTCGTGGGGCTCTTCGGGGTGTGTGACCC 3900
Qy 3901 GGGGGTCCGAAGGGGTGACTTCATACCGTTGAGTCTATGGAACCTACATGCGGT 3960
Dd 3901 GGGGGTCCGAAGGGGTGAGTTCATACCGTTGAGTCTATGGAACCTACATGCGGT 3960
Qy 3961 CTCGGTCTTTCAGACAACTAAACCCCGCGCTGTACCGCAGATTTCCAAAGTGGAC 4020
Dd 3961 CTCGGTCTTTCAGACAACTAAACCCCGCGCTGTACCGCAGATTTCCAAAGTGGAC 4020
Qy 4021 ATCTGACGCTCTTACTGCGAGCGGCAAGACCAAAAGTGGCGGTGCTATGACGCC 4080
Dd 4021 ATCTGACGCTCTTACTGCGAGCGGCAAGACCAAAAGTGGCGGTGCTATGACGCC 4080

Qy 4081 AAGGTTCAAGGTGCTCGTCTGTAAACCGTCCGTTCGCGCACCTTAGGGTTTGGGGGT 4140
Dd 4081 AAGGTTCAAGGTGCTCGTCTGTAAACCGTCCGTTCGCGCACCTTAGGGTTTGGGGGT 4140
Qy 4141 ATATGTTCAAGGCACAGGTATCGACCTTAACATCAGATCGGGTGAAGCACTTACCA 4200
Dd 4141 ATATGTTCAAGGCACAGGTATCGACCTTAACATCAGATCGGGTGAAGCACTTACCA 4200
Qy 4201 CGGCGGCTCCATTACGTACTCCACCTATGGCAAGTTCTTGGCGACGGTGGCTTCTG 4260
Dd 4201 CGGCGGCTCCATTACGTACTCCACCTATGGCAAGTTCTTGGCGACGGTGGCTTCTG 4260
Qy 4261 GGGCGGCTATGACATCAATATGATGAGTGCCTCAACTGACTCGACTACCATCT 4320
Dd 4261 GGGCGGCTATGACATCAATATGATGAGTGCCTCAACTGACTCGACTACCATCT 4320
Qy 4321 TGGGCATCGGCACAGTCTTGACCAAGGGAGAGCGGTGAGAGCGGCTCGTGTGCTG 4380
Dd 4321 TGGGCATCGGCACAGTCTTGACCAAGGGAGAGCGGTGAGAGCGGCTCGTGTGCTG 4380
Qy 4381 CCACGCTTACACCTCCGGATCGGTTACCGTGCACACCCCAATATCGAGGAAATAGGCC 4440
Dd 4381 CCACGCTTACACCTCCGGATCGGTTACCGTGCACACCCCAATATCGAGGAAATAGGCC 4440
Qy 4441 TGTCCAAACATGGAGAGATCCCTTCTATGGCAAGGCCATCCCATTTAGGCCCATCAAG 4500
Dd 4441 TGTCCAAACATGGAGAGATCCCTTCTATGGCAAGGCCATCCCATTTAGGCCCATCAAG 4500
Qy 4501 GGGGAGGCTCTCATTTTCTGCGCATTTCCAAAGAAATGTGACGAGTCCGCGCAAGC 4560
Dd 4501 GGGGAGGCTCTCATTTTCTGCGCATTTCCAAAGAAATGTGACGAGTCCGCGCAAGC 4560
Qy 4561 TGACAGGCTCGGACTGAAAGCTGTAGCATATTAACGGGGCTTGTATGTCCGCTCATAC 4620
Dd 4561 TGACAGGCTCGGACTGAAAGCTGTAGCATATTAACGGGGCTTGTATGTCCGCTCATAC 4620
Qy 4621 CGCTATCGGAGAGCTCGTGTGTCGCAACAGACGCTCTAATGACGGGTTTACCGGG 4680
Dd 4621 CGCTATCGGAGAGCTCGTGTGTCGCAACAGACGCTCTAATGACGGGTTTACCGGG 4680
Qy 4681 ATTTGACTGATGATGCAATACATGTGTACCCAGACAGTCTGACTTCACTTGG 4740
Dd 4681 ATTTGACTGATGATGCAATACATGTGTACCCAGACAGTCTGACTTCACTTGG 4740
Qy 4741 ATCCACCTTCAACATTTGAGACGACCGTCCCGCAAGACGCGGTGCGCTCGCAAC 4800
Dd 4741 ATCCACCTTCAACATTTGAGACGACCGTCCCGCAAGACGCGGTGCGCTCGCAAC 4800
Qy 4801 GCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC 4860
Dd 4801 GCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC 4860
Qy 4861 GGGCTCGGGCATGTTGCAATTTCTGCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
Dd 4861 GGGCTCGGGCATGTTGCAATTTCTGCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
Qy 4921 GGTATGACTCAGCGCCCTGAGACCTCGGTTAGGTTTCGGGCTTACCTAATACACAG 4980
Dd 4921 GGTATGACTCAGCGCCCTGAGACCTCGGTTAGGTTTCGGGCTTACCTAATACACAG 4980
Qy 4981 GGTTCGCCGTCTGCCAGGACCATCTGGAGTTCTGGAGAGCGTCTTACAGGCTTCAACC 5040
Dd 4981 GGTTCGCCGTCTGCCAGGACCATCTGGAGTTCTGGAGAGCGTCTTACAGGCTTCAACC 5040
Qy 5041 ACATAGATGCCACTTCTCTGCCAGACTAAACAGGAGGAGCAACTTCTTCTTACCTGG 5100
Dd 5041 ACATAGATGCCACTTCTCTGCCAGACTAAACAGGAGGAGCAACTTCTTCTTACCTGG 5100
Qy 5101 TGGCATATCAAGCTACAGTGTGCGCGGCTCAAGCTCCACCTCCATCTGTTGGGACCAA 5160
Dd 5101 TGGCATATCAAGCTACAGTGTGCGCGGCTCAAGCTCCACCTCCATCTGTTGGGACCAA 5160

QY	5161	TGTGGAAGTGTCTCATACGGCTGAACCTTACACTGCACGGGCAACACCCCTGCTGTATA	5220
DB	5161		5220
QY	5221	TGTGGAAGTGTCTCATACGGCTGAACCTTACACTGCACGGGCAACACCCCTGCTGTATA	5220
DB	5221		5220
QY	5281	GGCTAGGAGCGGTCCAAAATGAGGTCATCCTCTACACACCCCATAAATACATCATGG	5280
DB	5221	GGCTAGGAGCGGTCCAAAATGAGGTCATCCTCACACCCCATAAATACATCATGG	5280
QY	5281	CATTGCATGTCGGCTGACCTGGAGGTCGTCACTAGCACCTTGGTGCTGTGTAGCGGAGTCC	5340
DB	5281	CATTGCATGTCGGCTGACCTGGAGGTCGTCTACTAGCACCTTGGTGCTGTGTAGCGGAGTCC	5340
QY	5341	TTGCAGCTTTGGCGGCATACTGCTCAGCAGGCGAGTGTGTCAATTGTGGGCAAGATCA	5400
DB	5341	TTGCAGCTTTGGCGGCATACTGCTCAGCAGGCGAGTGTGTCAATTGTGGGCGAGATCA	5400
QY	5401	TCTTTGTCGGGAAGCCGAGTGTCTTCCGACAGGGAAGTCTCTACACAGGAGTTCCGATG	5460
DB	5401	TCTTTGTCGGGAAGCCAGCTGTCTTTCCGACAGGGAAGTCTCTTACCAGGAGTTCCGATG	5460
QY	5461	AGATGGAAGAGTGTGCTCTCAAACTTCTTACATCGAGCAGGGAATGCAGCTCGCCGAGC	5520
DB	5461	AGATGGAAGAGTGTGCTCTCAAACTTCTTACATCGAGCAGGGAATGCAGCTCGCCGAGC	5520
QY	5521	AATTCAAGAAAAAGCGCTCGGGTTGTTCGAAACGGCCACCAAGCAAGCGGAGGTGCTGTG	5580
DB	5521	AATTCAAGAAAAAGCGCTCGGGTTGTTCGAAACGGCCACCAAGCAAGCGGAGGTGCTGTG	5580
QY	5581	CTCCGCTGTGTGAGTTCGAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCATGTGGA	5640
DB	5581	CTCCGCTGTGTGAGTTCGAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCATGTGGA	5640
QY	5641	ATTTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCCTTGGAAAAACCCCGCA	5700
DB	5641	ATTTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCCTTGGAAAAACCCCGCA	5700
QY	5701	TAGCATCATTTGATGGCATTTACAGCTTCTATCACTAGCCCGCTCAACACCCAAAAACACC	5760
DB	5701	TAGCATCATTTGATGGCATTTACAGCTTCTATCACTAGCCCGCTCAACACCCAAAAACACC	5760
QY	5761	TCCTGTTTAAACATCTTGGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGCGT	5820
DB	5761	TCCTGTTTAAACATCTTGGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGCGT	5820
QY	5821	CAGCTTTCTGGGCGCCGGCATCCCGAGCGGCTGTGTGGCAGCATAGGCTTTGGGAAGG	5880
DB	5821	CAGCTTTCTGGGCGCCGGCATCCCGAGCGGCTGTGTGGCAGCATAGGCTTTGGGAAGG	5880
QY	5881	TGCTCGTGAGCATCTTGGCGGGCTATGGGGCAGGGGTAGCGGGCAGCTCGTGGCCCTTTA	5940
DB	5881	TGCTCGTGAGCATCTTGGCGGGCTATGGGGCAGGGGTAGCGGGCAGCTCGTGGCCCTTTA	5940
QY	5941	AGGTCATAGCGGGCAGGTGCCCTCCACCGAGGACCTGTGTCACTTACTCCCTGCCATCC	6000
DB	5941	AGGTCATAGCGGGCAGGTGCCCTCCACCGAGGACCTGTGTCACTTACTCCCTGCCATCC	6000
QY	6001	TCTCTCTGTGTGCCCTGTGTGTGGGGTGTGTGCGCAGCAATACTGCGTCGGCAGCTGG	6060
DB	6001	TCTCTCTGTGTGCCCTGTGTGTGGGGTGTGTGCGCAGCAATACTGCGTCGGCAGCTGG	6060
QY	6061	GCCCGGGAGAGGGGCTGTGCACTGGATGAACCGGCTGATAGCGTTGCTTTCGGGGGTA	6120
DB	6061	GCCCGGGAGAGGGGCTGTGCACTGGATGAACCGGCTGATAGCGTTGCTTTCGGGGGTA	6120
QY	6121	ACCAAGTCTCCCTACGCACTATGTGCCTTGAGAGGAGCGCTGCAGCAGTGTCACTCAGA	6180
DB	6121	ACCAAGTCTCCCTACGCACTATGTGCCTTGAGAGGAGCGCTGCAGCAGTGTCACTCAGA	6180
QY	6181	TCCTCTCTAGCCTTACCATCACTCAACTGTCTGAAGCGGCTCCACAGTGGATTAATGAGG	6240
DB	6181	TCCTCTCTAGCCTTACCATCACTCAACTGTCTGAAGCGGCTCCACAGTGGATTAATGAGG	6240
QY	6241	ACTGCTCTACGCCATGTCTCCGGCTCGTGGCTAAGGGATGTTTGGGATTTGATATGCAAGG	6300

Db 7321 CACGAGAGAAAGGACGGTGTCTGTACACAGAATCCAATGTGTCTTCTGCCTTGGCGGAGC 7380
Qy TCGCCACTAAGACCTTTCGGTAGCTCCGGATCGTCCGCCGTTGTAGACGGCACGGGACCG 7440
Db TCGCCACTAAGACCTTTCGGTAGCTCCGGATCGTCCGCCGTTGTAGACGGCACGGGACCG 7440
Qy CCCTTCTGTGACTCGCTCCGACGACGGTGAACAAAGSATCCGACGTTGAGTCTGTA 7500
Db CCCTTCTGTGACTCGCTCCGACGACGGTGAACAAAGSATCCGACGTTGAGTCTGTA 7500
Qy CCATGCCCCCTTGAAGGGGAGCCGGGGAGCCCGGATCTCAGCGACGGGTCTTGGTCTA 7560
Db CCATGCCCCCTTGAAGGGGAGCCGGGGAGCCCGGATCTCAGCGACGGGTCTTGGTCTA 7560
Qy CCGTGAGTGAGAGGCTAGTGAGGATGTCGTCTGCTCAATGTCCTATACGTGGAAG 7620
Db CCGTGAGTGAGAGGCTAGTGAGGATGTCGTCTGCTCAATGTCCTATACGTGGAAG 7620
Qy GCGCCCTGATCACGCCATGCGCTGCGGAGGAAAGTAAGCTGCCATCAACCCGTTGAGCA 7680
Db GCGCCCTGATCACGCCATGCGCTGCGGAGGAAAGTAAGCTGCCATCAACCCGTTGAGCA 7680
Qy ACTCTTTGCTGCTCACCAACATGCTACGCCACAACATPCCCGAGCGCAAGCCTCC 7740
Db ACTCTTTGCTGCTCACCAACATGCTACGCCACAACATPCCCGAGCGCAAGCCTCC 7740
Qy GGCAGAGAGGTACCTTTGACAGATTCGAGTTCCTCGATGATCATTTACCGGGACGTAC 7800
Db GGCAGAGAGGTACCTTTGACAGATTCGAGTTCCTCGATGATCATTTACCGGGACGTAC 7800
Qy TCAAGGAGATGAAGCGAAGCGCTCCACAGTTAAGGCTTAAGCTTCTATATAGAGGAGG 7860
Db TCAAGGAGATGAAGCGAAGCGCTCCACAGTTAAGGCTTAAGCTTCTATATAGAGGAGG 7860
Qy CTGCAAGGTGAGCGCCCGCACATTCGGGCAAAATCCAAATTTGGCTATCGGGCAAAAGGACG 7920
Db CTGCAAGGTGAGCGCCCGCACATTCGGGCAAAATCCAAATTTGGCTATCGGGCAAAAGGACG 7920
Qy TCCGGAACCTATCCAGCAGGCGGTTAACCATCCGCTCCGTTGGGAGGACTTGTGG 7980
Db TCCGGAACCTATCCAGCAGGCGGTTAACCATCCGCTCCGTTGGGAGGACTTGTGG 7980
Qy AAGACACTGAACACCAATTTGACACACCATCATGGCAAAAAGTAGGTTTTCTCGCTCC 8040
Db AAGACACTGAACACCAATTTGACACACCATCATGGCAAAAAGTAGGTTTTCTCGCTCC 8040
Qy AACGAGAGAGGAGCGCGCAAGCAGCTCGCTTTATCGTATPCCAGACCTGGGAGTTC 8100
Db AACGAGAGAGGAGCGCGCAAGCAGCTCGCTTTATCGTATPCCAGACCTGGGAGTTC 8100
Qy GTGTATGCGAGAGATGGCCCTTTACGAGCTGCTCCACCTTCTCAGGCGGTGATGG 8160
Db GTGTATGCGAGAGATGGCCCTTTACGAGCTGCTCCACCTTCTCAGGCGGTGATGG 8160
Qy GCTCTCTATACGGAATTCAAATACTCCCCAAGCAGCGGTCCAGTTCTGTTGAATACCT 8220
Db GCTCTCTATACGGAATTCAAATACTCCCCAAGCAGCGGTCCAGTTCTGTTGAATACCT 8220
Qy GGAATCAAAAGAAATGCCCTATGGCTTCTCATATGACACCCGCTGTTTGACTCAACGG 8280
Db GGAATCAAAAGAAATGCCCTATGGCTTCTCATATGACACCCGCTGTTTGACTCAACGG 8280
Qy TCACCTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACCAATGTTGATCTGGCCCCGG 8340
Db TCACCTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACCAATGTTGATCTGGCCCCGG 8340
Qy AGGCCAGACAGGCCATAAGGTCCGCTCACAGACGGCTTTACATCGGGGTCCCTTGACTA 8400
Db AGGCCAGACAGGCCATAAGGTCCGCTCACAGACGGCTTTACATCGGGGTCCCTTGACTA 8400
Qy ACTCAAAAGGGCAGAACTGCGGTTATCGCCGGTGCAGCGCAAGTGCCTGACGACTA 8460
Db ACTCAAAAGGGCAGAACTGCGGTTATCGCCGGTGCAGCGCAAGTGCCTGACGACTA 8460

Qy 8461 GCTCGGCTAATACCTCACATGTTACTTTGAAGGCCACTGACGCCCTGTGAGCTGCAAAAGC 8520
Db 8461 GCTCGGCTAATACCTCACATGTTACTTTGAAGGCCACTGACGCCCTGTGAGCTGCAAAAGC 8520
Qy TCCAGGACTGCAATGCTCGTGAAACGGAGACGACCTTGTCTGTTATCTGTGAAAGCGGG 8580
Db TCCAGGACTGCAATGCTCGTGAAACGGAGACGACCTTGTCTGTTATCTGTGAAAGCGGG 8580
Qy GAAACCCAGGAGGATGCGCGGCCCTCAGAGCTTTCACGGAGGCTATGACTAGGTATTCGG 8640
Db GAAACCCAGGAGGATGCGCGGCCCTCAGAGCTTTCACGGAGGCTATGACTAGGTATTCGG 8640
Qy CCCCCCGGGGATCGGCCCAACAGAAATACGACCTGGAGCTGATAACATCATCTTCT 8700
Db CCCCCCGGGGATCGGCCCAACAGAAATACGACCTGGAGCTGATAACATCATCTTCT 8700
Qy CCAATGTGTCACTGCGCGACGATGATCTGGCAAAAGGTTATCTACCTCAACCCGTGACC 8760
Db CCAATGTGTCACTGCGCGACGATGATCTGGCAAAAGGTTATCTACCTCAACCCGTGACC 8760
Qy CCACACCCCTTTCAGCGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAACTCTT 8820
Db CCACACCCCTTTCAGCGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAACTCTT 8820
Qy GGCTAGGCAATATCATGATGATCGGCCACCTTATGGCAAGGATGATTTCTGATGACTC 8880
Db GGCTAGGCAATATCATGATGATCGGCCACCTTATGGCAAGGATGATTTCTGATGACTC 8880
Qy ACTTTTCTCCATCTTCTAGCTCAAGAGCAACTTTGAAAAAGCCCTGGATTTGCAGATCT 8940
Db ACTTTTCTCCATCTTCTAGCTCAAGAGCAACTTTGAAAAAGCCCTGGATTTGCAGATCT 8940
Qy ACGGGCTTGTACTCCATTCAGCCACTTGACTACCTCAGATCATTTGAAACGACTCCATG 9000
Db ACGGGCTTGTACTCCATTCAGCCACTTGACTACCTCAGATCATTTGAAACGACTCCATG 9000
Qy GTCTTAGGGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCAT 9060
Db GTCTTAGGGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCAT 9060
Qy GCCTCAGSAAACTTTGGGTACCCCTTTGCAACCTCGAGACATCGGGCCAGAACTGTC 9120
Db GCCTCAGSAAACTTTGGGTACCCCTTTGCGAACTCGGAGACATCGGGCCAGAACTGTC 9120
Qy GCGCTAAGCTACTGTCAGGGGGGAGGGCGGCACTTGTGGCAGATACCTCTTTAACT 9180
Db GCGCTAAGCTACTGTCAGGGGGGAGGGCGGCACTTGTGGCAGATACCTCTTTAACT 9180
Qy GGGCAGTAAGGACCAAGCTTAACTCACTCAATCCCGCCGCGTCCAGCTGGACTTGT 9240
Db GGGCAGTAAGGACCAAGCTTAACTCACTCAATCCCGCCGCGTCCAGCTGGACTTGT 9240
Qy CTGCTGCTGCTCGCTGCTTACAGCGGGGAGACATATATCACAGCTGTCTCGTGCCC 9300
Db CTGCTGCTGCTCGCTGCTTACAGCGGGGAGACATATATCACAGCTGTCTCGTGCCC 9300
Qy GACCCCGCTGCTTCCGTTGCTCTACTCTTCTGTTAGGGTAGGCATTTACCTGCG 9360
Db GACCCCGCTGCTTCCGTTGCTCTACTCTTCTGTTAGGGTAGGCATTTACCTGCG 9360
Qy TCCCAACCCGATGAACGGGAGCTAAACCACTCCAGGCTTAAAGCCATTTCTGTTTTTTT 9420
Db TCCCAACCCGATGAACGGGAGCTAAACCACTCCAGGCTTAAAGCCATTTCTGTTTTTTT 9420
Qy TT 9480
Db TT 9480
Qy TTTTTCCCTCTTTAAAGTGCTGCTCCATCTTAGCCCTAGTACCGCTAGCTGTGAAGGT 9540
Db TTTTTCCCTCTTTAAAGTGCTGCTCCATCTTAGCCCTAGTACCGCTAGCTGTGAAGGT 9540

1441 ACTGGCTAAGGTTCTGATTGTGGCGCTACTCTTTGCGGCGTTTGACGGGAGACCCACA 1500
1441 ACTGGGCTAAGGTTCTGATTGTGGCGCTACTCTTTGCGGCGTTTGACGGGAGACCCACA 1500
1501 CGACGGGAGGGTGCGCGGCCACACACACTCCGGGTTCACTGCTTTTCTCATCTGGGG 1560
1501 CGACGGGAGGGTGCGCGGCCACACACACTCCGGGTTCACTGCTTTTCTCATCTGGGG 1560
1561 GGTCTCAGAAATCAGCTTGTGAATACCAACGGCAGCTGGCACATCAACAGGACTGCC 1620
1561 GGTCTCAGAAATCAGCTTGTGAATACCAACGGCAGCTGGCACATCAACAGGACTGCC 1620
1621 TAAATTTGAATGACTCCCTCCAAACTGGGTTCTTTGCGGCGTGTGTTTACGCACACAAGT 1680
1621 TAAATTTGAATGACTCCCTCCAAACTGGGTTCTTTGCGGCGTGTGTTTACGCACACAAGT 1680
1681 TCAACTCGTCCGGGTGCCGGAGCGCATGGCCAGCTGCCGCCCATTTGACTGGTTGCGCC 1740
1681 TCAACTCGTCCGGGTGCCGGAGCGCATGGCCAGCTGCCGCCCATTTGACTGGTTGCGCC 1740
1741 AGGGTGGGCCCCCATCACTATACCTAAGCCTAAACAGCTCGSATCAGAGGCCCTTATTGCT 1800
1741 AGGGTGGGCCCCCATCACTATACCTAAGCCTAAACAGCTCGSATCAGAGGCCCTTATTGCT 1800
1801 GGCATTACGCGCTCGACCGTGTGTCTGTACCCGCTCGCAGGTGTGTGTCAGTGT 1860
1801 GGCATTACGCGCTCGACCGTGTGTCTGTACCCGCTCGCAGGTGTGTGTCAGTGT 1860
1861 ATTGTTTACCCCAAGCCCTGTTGTGGTGGGACCAACGATCGTTCCGGTGTCCCTACGT 1920
1861 ATTGTTTACCCCAAGCCCTGTTGTGGTGGGACCAACGATCGTTCCGGTGTCCCTACGT 1920
1921 ATAGCTGGGGGAGAAATGACAGACAGTGTATCTCTCMAACACAGCGTCCGCCACAAAG 1980
1921 ATAGCTGGGGGAGAAATGACAGACAGTGTATCTCTCMAACACAGCGTCCGCCACAAAG 1980
1981 GCAACTGGTTCGGCTGTACATGGATGAATAGTACTGGGTTCACTAAGAGTCCGGAGGTC 2040
1981 GCAACTGGTTCGGCTGTACATGGATGAATAGTACTGGGTTCACTAAGAGTCCGGAGGTC 2040
2041 CCCCGTTAAACATCGGGGGGTCGGTAAACCGCACCTTGATCTGCCACCGACTGCTTCC 2100
2041 CCCCGTTAAACATCGGGGGGTCGGTAAACCGCACCTTGATCTGCCACCGACTGCTTCC 2100
2101 GGAAGACCCCCAGAGCTACTTACAAATATGTGGCTCGGGGCCCTGGTTGACACCTAGGT 2160
2101 GGAAGACCCCCAGAGCTACTTACAAATATGTGGCTCGGGGCCCTGGTTGACACCTAGGT 2160
2161 GCCTAGTACCTACCATACAGCTTTGGCCTACCCCTGCACTCTCAATTTTTCATCT 2220
2161 GCCTAGTACCTACCATACAGCTTTGGCCTACCCCTGCACTCTCAATTTTTCATCT 2220
2221 TTAAGGTTAGGATGTATGTGGGGGCGTGGAGCACAGGCTCAATGCGCATGCAATTGGA 2280
2221 TTAAGGTTAGGATGTATGTGGGGGCGTGGAGCACAGGCTCAATGCGCATGCAATTGGA 2280
2281 CTCGAGGAGAGCGCTGTAACTTTGGAGGACAGGGATAGGTCAGAACTCAGCCCGCTGCTGC 2340
2281 CTCGAGGAGAGCGCTGTAACTTTGGAGGACAGGGATAGGTCAGAACTCAGCCCGCTGCTGC 2340
2341 TGTCTACACAGAGTGGCAGATCTGCCCTGTGCTTTTACACCCCTACCGGCTTTATCCA 2400
2341 TGTCTACACAGAGTGGCAGATCTGCCCTGTGCTTTTACACCCCTACCGGCTTTATCCA 2400
2401 CTGGTTTGTATCCATCTCCATCAGAAACATCGTGGACGTGCAATACCTGTACGGGTAGGGT 2460
2401 CTGGTTTGTATCCATCTCCATCAGAAACATCGTGGACGTGCAATACCTGTACGGGTAGGGT 2460
2461 CAGCGTTTGTCTCTTTTGAATCAAAATGGGAGTACATCTCTGTTGCTTTTCTCTCTGG 2520
2461 CAGCGTTTGTCTCTTTTGAATCAAAATGGGAGTACATCTCTGTTGCTTTTCTCTCTGG 2520

2521 CAGACGCGCGGTGTGTGCTGCTTGTGATGATGCTGCTGATAGCCAGGCTGAGGCG 2580
2521 CAGACGCGCGGTGTGTGCTGCTTGTGATGATGCTGCTGATAGCCAGGCTGAGGCG 2580
2581 CTTTAGAGAACTTGGTGTGCTCAATGCGGCGTCCGTGGCGGAGCGGATGATTTCTCT 2640
2581 CTTTAGAGAACTTGGTGTGCTCAATGCGGCGTCCGTGGCGGAGCGGATGATTTCTCT 2640
2641 CTTTCTTGTGCTTCTTCTGCGCGCTGCTTAAAGGAGGCTGGCTCTCTGGGG 2700
2641 CTTTCTTGTGCTTCTTCTGCGCGCTGCTTAAAGGAGGAGGCTGGCTCTCTGGGG 2700
2701 CGTATGCTTTTATGGCGTATGGCGCTGCTCTGCTCTCTACTGCGTTTACCAACAG 2760
2701 CGTATGCTTTTATGGCGTATGGCGCTGCTCTGCTCTCTACTGCGTTTACCAACAG 2760
2761 CTTAGCCCTTGGACCGGAGATGGTGCATCGTGGGGGGTGGCTTCTTGTAGTCTGG 2820
2761 CTTAGCCCTTGGACCGGAGATGGTGCATCGTGGGGGGTGGCTTCTTGTAGTCTGG 2820
2821 TATTCTTGACCTTGTCAACATACCAAGTGTCTTCTCACTAGGCTCATATGGTGTAC 2880
2821 TATTCTTGACCTTGTCAACATACCAAGTGTCTTCTCACTAGGCTCATATGGTGTAC 2880
2881 AATACTTTATCACCAGAGCGGACATGCAAGTGTGGTCCCGCTCAACGTTT 2940
2881 AATACTTTATCACCAGAGCGGACATGCAAGTGTGGTCCCGCTCAACGTTT 2940
2941 GGGAGGCGCGGATGCCATCATCTCTCCTCAGTGTGGGTTTCACTCAGAGTTAAATTTTG 3000
2941 GGGAGGCGCGGATGCCATCATCTCTCCTCAGTGTGGGTTTCACTCAGAGTTAAATTTTG 3000
3001 ACATCACCAGAACTCTGCTCGCCATACCTCGGCCGCTCATGTGCTCCAGGCTGCATAA 3060
3001 ACATCACCAGAACTCTGCTCGCCATACCTCGGCCGCTCATGTGCTCCAGGCTGCATAA 3060
3061 CGAGAGTCCGCTACTTCTGCGCGCTCAAGGCTCATTCGTGCATGCTAGTTCGCGAA 3120
3061 CGAGAGTCCGCTACTTCTGCGCGCTCAAGGCTCATTCGTGCATGCTAGTTCGCGAA 3120
3121 AAGTCCGCGGGGTCATTATGTCCAAATGGTCTTCAAGTGTGGGCGCTGACAGGTA 3180
3121 AAGTCCGCGGGGTCATTATGTCCAAATGGTCTTCAAGTGTGGGCGCTGACAGGTA 3180
3181 CGTAGGTTTATAACCATCTTACCCACCTGCGGGAATGCGGCGCTACGAGACC 3240
3181 CGTAGGTTTATAACCATCTTACCCACCTGCGGGAATGCGGCGCTACGAGACC 3240
3241 TTGCGGTGGCGGTAGAGCCGCTGCTTCTCCGCCATGGAGACCAAGGTCATCACCTGGG 3300
3241 TTGCGGTGGCGGTAGAGCCGCTGCTTCTCCGCCATGGAGACCAAGGTCATCACCTGGG 3300
3301 GAGCAGACACCGCTCGTGTGGGACATCATCTTGGGCTACCCGCTCTCCGCCGAGGG 3360
3301 GAGCAGACACCGCTCGTGTGGGACATCATCTTGGGCTACCCGCTCTCCGCCGAGGG 3360
3361 GGAAGGAGATATTTTGGGACCGGCTGATGCTCGAAGGGCAAGGTTGGGACTCTTTG 3420
3361 GGAAGGAGATATTTTGGGACCGGCTGATGCTCGAAGGGCAAGGTTGGGACTCTTTG 3420
3421 CGCCCATCAGCGCTTACTCCCAACAAACGCGGGGCTACTTGGTGCATCATCATAGCC 3480
3421 CGCCCATCAGCGCTTACTCCCAACAAACGCGGGGCTACTTGGTGCATCATCATAGCC 3480
3481 TCACAGGCGGACCAAGAACCAAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAAC 3540
3481 TCACAGGCGGACCAAGAACCAAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAAC 3540
3541 AATCTTTCTGGCGACTTCATCAACGGGCTGTGGACTCTCTACCACTGGCTGGCT 3600
3541 AATCTTTCTGGCGACTTCATCAACGGGCTGTGGACTCTCTACCACTGGCTGGCT 3600
3601 CGAAGACCTTAGCCGCTCCAAAGAGTCCAATCACCCAAATGTACACCAATGTAGACCTGG 3660

Db 3601 CGAAGACCTTAGCCGGTCCAAAGGTCCAATACCCAAATGTACACCAATGTAGACCTGG 3660
Qy 3661 ACCTCGTGGGTGACGCGCCCCCGGGGCGCTCCATGACACCAATGACAGTGTGGCA 3720
Db 3661 ACCTCGTGGGTGACGCGCCCCCGGGGCGCTCCATGACACCAATGACAGTGTGGCA 3720
Qy 3721 GCTCGGACCTTTACTTGGTACAGAGACANTGTGATGTATTCGGGTGCGCGGAGGCG 3780
Db 3721 GCTCGGACCTTTACTTGGTACAGAGACANTGTGATGTATTCGGGTGCGCGGAGGCG 3780
Qy 3781 ACAGCAGGGGAAGTCTACTCTCCCCCAGGCGCGTCTCTACCTCAAGAGCTCCTCGGGTG 3840
Db 3781 ACAGCAGGGGAAGTCTACTCTCCCCCAGGCGCGTCTCTACCTCAAGAGCTCCTCGGGTG 3840
Qy 3841 GTCCATTGCTTTGGCCCTTGGGGCAGCGTGGGGCGCTTCCGGGCTGCTGTGTGCAACC 3900
Db 3841 GTCCATTGCTTTGGCCCTTGGGGCAGCGTGGGGCGCTTCCGGGCTGCTGTGTGCAACC 3900
Qy 3901 GGGGGTCCGGAAGCGGTGGACTTTCATACCCGTTGAGTCTATGGAACATACCAATGCGGT 3960
Db 3901 GGGGGTCCGGAAGCGGTGGACTTTCATACCCGTTGAGTCTATGGAACATACCAATGCGGT 3960
Qy 3961 CTCGGTCTTACAGACAACCTCAACCCCGCGGCTGTACCGCAGACATTCCAAGTGGCAC 4020
Db 3961 CTCGGTCTTACAGACAACCTCAACCCCGCGGCTGTACCGCAGACATTCCAAGTGGCAC 4020
Qy 4021 ATCTGCAAGCTCTACTTGGCAGCGGCAAGACCAAAAGTCCCGCTATGCAAGGCC 4080
Db 4021 ATCTGCAAGCTCTACTTGGCAGCGGCAAGACCAAAAGTCCCGCTATGCAAGGCC 4080
Qy 4081 AAGGGTCAAGGGTCTCGTCTGAAACCCGTCGTTGGCGCACCTTAGGGTTTGGGGGT 4140
Db 4081 AAGGGTCAAGGGTCTCGTCTGAAACCCGTCGTTGGCGCACCTTAGGGTTTGGGGGT 4140
Qy 4141 ATATGTCGAAGGCACACGGTATGACCCCTACATCAGNACTGGGGTAAGNACCAATTACCA 4200
Db 4141 ATATGTCGAAGGCACACGGTATGACCCCTACATCAGNACTGGGGTAAGNACCAATTACCA 4200
Qy 4201 CGGCGGCTCCATTACGTPACTCCACCTATGCAAGTTCTCTTGGCGACGGTGGCTGTCTG 4260
Db 4201 CGGCGGCTCCATTACGTPACTCCACCTATGCAAGTTCTCTTGGCGACGGTGGCTGTCTG 4260
Qy 4261 GGGGCGCTATGACATCATATATGTGATGAGTGCCACTCACTGACTCGACTACCATCT 4320
Db 4261 GGGGCGCTATGACATCATATATGTGATGAGTGCCACTCACTGACTCGACTACCATCT 4320
Qy 4321 TGGGCATCGGCACAGTCTCGACCAAGCGGAGACGGCTGGAGCGCGCTCGTGGCTCG 4380
Db 4321 TGGGCATCGGCACAGTCTCGACCAAGCGGAGACGGCTGGAGCGCGCTCGTGGCTCG 4380
Qy 4381 CCACGGCTACACCTCCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAATAGGCC 4440
Db 4381 CCACGGCTACACCTCCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAATAGGCC 4440
Qy 4441 TGTCCAAATGAGAGATCCCTTCTATGCAAGACCATCCCAATTGAGGCCATCAAGG 4500
Db 4441 TGTCCAAATGAGAGATCCCTTCTATGCAAGACCATCCCAATTGAGGCCATCAAGG 4500
Qy 4501 GGGGAGGCATCTCATTTTCTGCCATTCGAAGAAATGTGACGAGCTCCCGCAAGC 4560
Db 4501 GGGGAGGCATCTCATTTTCTGCCATTCGAAGAAATGTGACGAGCTCCCGCAAGC 4560
Qy 4561 TGACAGGCTCGGACTGAACCGTGTAGCATATTACCGGGGCTTGATGTGCGTCTATC 4620
Db 4561 TGACAGGCTCGGACTGAACCGTGTAGCATATTACCGGGGCTTGATGTGCGTCTATC 4620
Qy 4621 CGCCTATCGGAGACGTGTGCTGGCAACAGACGCTCTAATGACGGTTTACCGGCG 4680
Db 4621 CGCCTATCGGAGACGTGTGCTGGCAACAGACGCTCTAATGACGGTTTACCGGCG 4680
Qy 4681 ATTTTGACTCAGTGCAGTGCATATGTGTCCACGACAGTGCAGTTCAGCTTGG 4740

Db 4681 ATTTTGACTCAGTGCAGTGCATATCATGTGTACCCAGACAGTGCAGCTTCAGCTTG 4740
Qy 4741 ATCCCACTTTACCAATTGAGACGACACCGTGGCCCCAAGACGGGTGTGCGCTCGCAAC 4800
Db 4741 ATCCCACTTTACCAATTGAGACGACACCGTGGCCCCAAGACGGGTGTGCGCTCGCAAC 4800
Qy 4801 GCGAGGTAGAACTGGCAGGGTAGGAGTGGCATCTACAGTTTGTGACTCCAGGAGAAC 4860
Db 4801 GCGAGGTAGAACTGGCAGGGTAGGAGTGGCATCTACAGTTTGTGACTCCAGGAGAAC 4860
Qy 4861 GGCCTCGGGCATGTTTCGATTCTTCGTCCTGTGTGAGTGTATGACGCGGGCTGTGCTT 4920
Db 4861 GGCCTCGGGCATGTTTCGATTCTTCGTCCTGTGTGAGTGTATGACGCGGGCTGTGCTT 4920
Qy 4921 GGTATGAGCTCACGCCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACTAAATACACAG 4980
Db 4921 GGTATGAGCTCACGCCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACTAAATACACAG 4980
Qy 4981 GGTGGCCGCTGCGCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGCTCACCC 5040
Db 4981 GGTGGCCGCTGCGCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGCTCACCC 5040
Qy 5041 ACATAGATGCCACTTCTCTGCCAGACTAAACAGGACGAGACAACTTTCCTTACCTGG 5100
Db 5041 ACATAGATGCCACTTCTCTGCCAGACTAAACAGGACGAGACAACTTTCCTTACCTGG 5100
Qy 5101 TGGCATATCAAGCTTACAGGCTGCGCCAGGGCTCAAGCTCCACCTCCATCGTGGACAAA 5160
Db 5101 TGGCATATCAAGCTTACAGGCTGCGCCAGGGCTCAAGCTCCACCTCCATCGTGGACAAA 5160
Qy 5161 TGTGGAAGTGTCTCATAGGCTGAAACCTACCTGACGCGGCCAACACCCCTGCTGTATA 5220
Db 5161 TGTGGAAGTGTCTCATAGGCTGAAACCTACCTGACGCGGCCAACACCCCTGCTGTATA 5220
Qy 5221 GGCTAGGAGCGCTCCAAAATGAGTCTATCTCACACCCCAATAAATACATCATGG 5280
Db 5221 GGCTAGGAGCGCTCCAAAATGAGTCTATCTCACACCCCAATAAATACATCATGG 5280
Qy 5281 CATGCATGTCCGGTGCCTCGAGGTCGTCACTAGCACCTGGGTCTGGTAGCGGAGTCC 5340
Db 5281 CATGCATGTCCGGTGCCTCGAGGTCGTCACTAGCACCTGGGTCTGGTAGCGGAGTCC 5340
Qy 5341 TTGCAGCTTTGGCGGCATATCTGCTGACGACAGGAGTGTGTCATTTGTGGGACAGGATCA 5400
Db 5341 TTGCAGCTTTGGCGGCATATCTGCTGACGACAGGAGTGTGTCATTTGTGGGACAGGATCA 5400
Qy 5401 TCTTGTCCGGAGACCGAGCTGTGTTCCGACAGGAGTCTCTACAGAGTTCGATG 5460
Db 5401 TCTTGTCCGGAGACCGAGCTGTGTTCCGACAGGAGTCTCTACAGAGTTCGATG 5460
Qy 5461 AGATGGAAGAGTGTGCTCAAACTTCTTACATCGAGCAGGGAATGAGCTCGCCGAGC 5520
Db 5461 AGATGGAAGAGTGTGCTCAAACTTCTTACATCGAGCAGGGAATGAGCTCGCCGAGC 5520
Qy 5521 AATTCAAGCAAAAGGCGCTCGGGTTGTGCAACCGGCCAACAGCAAGCGGAGCTGCTG 5580
Db 5521 AATTCAAGCAAAAGGCGCTCGGGTTGTGCAACCGGCCAACAGCAAGCGGAGCTGCTG 5580
Qy 5581 CTCCCGTGTGGAGTCAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCACAATGTGA 5640
Db 5581 CTCCCGTGTGGAGTCAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCACAATGTGA 5640
Qy 5641 ATTTTCATCAGCGGAATACAGTACCTAGCAGCTTATCCACTCTCCCTGGAAACCCCGGA 5700
Db 5641 ATTTTCATCAGCGGAATACAGTACCTAGCAGCTTATCCACTCTCCCTGGAAACCCCGGA 5700
Qy 5701 TAGCATCATTTGATGGCATTTTACAGCTTCTATCACTAGCCGCTCACCAACCAACACCC 5760
Db 5701 TAGCATCATTTGATGGCATTTTACAGCTTCTATCACTAGCCGCTCACCAACCAACACCC 5760
Qy 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGGGT 5820
Db 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGGGT 5820

QY	5821	CAGCTTTTCGTGGCGCGCGCATCGCGGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG	5880
DB	5821	CAGCTTTTCGTGGCGCGCGCATCGCGGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG	5880
QY	5881	TGCTCGTGAGACATCTTGGCGGGCTATGGGGCAGGGGTAGCGGCGACACTCGTGGGCTTTTA	5940
DB	5881	TGCTCGTGAGACATCTTGGCGGGCTATGGGGCAGGGGTAGCGGCGACACTCGTGGGCTTTTA	5940
QY	5941	AGGTATGAGCGCGGAGAGTGCCTCGA CCGAGGACCTGGTCAACTTACTTCCCTGCATCC	6000
DB	5941	AGGTATGAGCGCGGAGAGTGCCTCGA CCGAGGAGCTGGTCAACTTACTTCCCTGCATCC	6000
QY	6001	TCCTCTCTGGTCCCTGGTTCGGGTCTGTGGCAGCAATACTTGGTGGCAGCTGG	6060
DB	6001	TCCTCTCTGGTCCCTGGTTCGGGTCTGTGGCAGCAATACTTGGTGGCAGCTGG	6060
QY	6061	GCCGGGAGAGGGGGTGTGCAGTGGATGAACCGGCTGATAGCTTTCGGTTCGCGGGGTA	6120
DB	6061	GCCGGGAGAGGGGGTGTGCAGTGGATGAACCGGCTGATAGCTTTCGGTTCGCGGGGTA	6120
QY	6121	ACCACTCTCCCTTAGCACTATGTGCTGAGAGGACGCTGCAGCACTGCTCAGA	6180
DB	6121	ACCACTCTCCCTTAGCACTATGTGCTGAGAGGACGCTGCAGCACTGCTCAGA	6180
QY	6181	TCCTCTTAGCCTTACCATCACTCAACTCTGAAGCGGCTCCACAGTGGATTAATGAGG	6240
DB	6181	TCCTCTTAGCCTTACCATCACTCAACTCTGAAGCGGCTCCACAGTGGATTAATGAGG	6240
QY	6241	ACTGTCTTACGCCATGCTCCGGCTCGTGCTTAAGGGATGTTTGGATTTGGATATGACGG	6300
DB	6241	ACTGTCTTACGCCATGCTCCGGCTCGTGCTTAAGGGATGTTTGGATTTGGATATGACGG	6300
QY	6301	TGTTGACTGACTTAAGACTCTGGCTCCAGTCCAAACTCTCTGCGGGGTTTACCGGAGTCC	6360
DB	6301	TGTTGACTGACTTAAGACTCTGGCTCCAGTCCAAACTCTCTGCGGGGTTTACCGGAGTCC	6360
QY	6361	CTTTCTCTGTCATGCCAAGCGGGTACAAGGAGTCTGGCGGGGAGCGCATCTGCATA	6420
DB	6361	CTTTCTCTGTCATGCCAAGCGGGTACAAGGAGTCTGGCGGGGAGCGCATCTGCATA	6420
QY	6421	CCACTTGCCTTACGAGACAGATCGCGGACATGTCAAAAACGGTTCCATGAGGATCG	6480
DB	6421	CCACTTGCCTTACGAGACAGATCGCGGACATGTCAAAAACGGTTCCATGAGGATCG	6480
QY	6481	TAGGGCTTAGAACCTGCAGCAACAGTGGCA CGGAACGTTCCCATCAACGCAACACCA	6540
DB	6481	TAGGGCTTAGAACCTGCAGCAACAGTGGCA CGGAACGTTCCCATCAACGCAACACCA	6540
QY	6541	CGGGACCTTGCACACCTCCCGGCGCCCACTATTCCAGGGGCTATGGCGGGTGGCTG	6600
DB	6541	CGGGACCTTGCACACCTCCCGGCGCCCACTATTTCAGGGGCGCTATGGCGGGTGGCTG	6600
QY	6601	CTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGCATGACCA	6660
DB	6601	CTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGCATGACCA	6660
QY	6661	CTGACAAAGTAAAGTCCCATGCTCCAGGTTCCGGGCCCGGAAATTTCTTCAAGGAGTGGATG	6720
DB	6661	CTGACAAAGTAAAGTCCCATGCTCCAGGTTCCGGGCCCGGAAATTTCTTCAAGGAGTGGATG	6720
QY	6721	GAGTGGGTTGCACAGGTACGCTCCGGCGTGCAACCTCTTCTACGGGAGGAGTCAAGT	6780
DB	6721	GAGTGGGTTGCACAGGTACGCTCCGGCGTGCAACCTCTTCTACGGGAGGAGTCAAGT	6780
QY	6781	TCCAGGTCCGGGCTCAACCAATACTTGGTCCGGGTCCGAGCTCCCATCCGAGCCCGAACCGG	6840
DB	6781	TCCAGGTCCGGGCTCAACCAATACTTGGTCCGGGTCCGAGCTCCCATCCGAGCCCGAACCGG	6840
QY	6841	ACGTAAACAGTGTCTTCTCATGCTCAACCGATCCCTCCACATTAACGAGAGACGGCTA	6900
DB	6841	ACGTAAACAGTGTCTTCTCATGCTCAACCGATCCCTCCACATTAACGAGAGACGGCTA	6900

QY	6901	AGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTTGT	6960
DB	6901	AGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTTGT	6960
QY	6961	CTGGGCTTCTTTGAGGCGACATGCTACCTACCCACCATGACTCCCCGAGCGTGAACCTCA	7020
DB	6961	CTGGGCTTCTTTGAGGCGACATGCTACCTACCCACCATGACTCCCCGAGCGTGAACCTCA	7020
QY	7021	TCGAGGCCAACCTCTTGTGGCGGAGGAGTGGGCGGAAACATCACTCCGCTGGAGTCAAG	7080
DB	7021	TCGAGGCCAACCTCTTGTGGCGGAGGAGTGGGCGGAAACATCACTCCGCTGGAGTCAAG	7080
QY	7081	AGAATAAGGTAGTAACTCTGGACTCTTTTCGAACCGCTTTCAGCGGAGGGGATGAGAGG	7140
DB	7081	AGAATAAGGTAGTAACTCTGGACTCTTTTCGAACCGCTTTCAGCGGAGGGGATGAGAGG	7140
QY	7141	AGATATCCGTTCGGGCGGAGATCTCTGCGAAATCCAGGAAATTCCTCAGGCTTGCCTCA	7200
DB	7141	AGATATCCGTTCGGGCGGAGATCTCTGCGAAATCCAGGAAATTCCTCAGGCTTGCCTCA	7200
QY	7201	TATGGGACGCGCGGACTACAAATCTCCACTGTAGAGTCTTGGAAAGACCCGGACTACG	7260
DB	7201	TATGGGACGCGCGGACTACAAATCTCCACTGTAGAGTCTTGGAAAGACCCGGACTACG	7260
QY	7261	TCCCTCCGGTGTACACCGGATGCCATTTGCCACCTTACCAAGGCTCTCTCAATACCACCTC	7320
DB	7261	TCCCTCCGGTGTACACCGGATGCCATTTGCCACCTTACCAAGGCTCTCTCAATACCACCTC	7320
QY	7321	CACGAGAAAGAGGACGTTGTCTGACAGAAATCCAAATGTGTCTTCTGCTTGGCGGAGC	7380
DB	7321	CACGAGAAAGAGGACGTTGTCTGACAGAAATCCAAATGTGTCTTCTGCTTGGCGGAGC	7380
QY	7381	TGCGCACTAAGACCTTCGGTAGCTCCGGATCGTGGCGGTTGATAGCGGACCGGACCG	7440
DB	7381	TGCGCACTAAGACCTTCGGTAGCTCCGGATCGTGGCGGTTGATAGCGGACCGGACCG	7440
QY	7441	CCCTTCTGACTGACCTGCCCTCCGACGCGTGACAAAGGATCCGACGTTGAGTCTGCTCT	7500
DB	7441	CCCTTCTGACTGACCTGCCCTCCGACGCGTGACAAAGGATCCGACGTTGAGTCTGCTCT	7500
QY	7501	CCATCCCCCTTGAAGGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTTA	7560
DB	7501	CCATCCCCCTTGAAGGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTTA	7560
QY	7561	CGTGAGTGAAGGAGCTAGTGAAGATGTCTGTCTGTCTCAATGTCTTATAGTGACAG	7620
DB	7561	CGTGAGTGAAGGAGCTAGTGAAGATGTCTGTCTGTCTCAATGTCTTATAGTGACAG	7620
QY	7621	GGGCTCTGATCAGGCATGCGCTGCGGAGGAAAGTAAAGCTGCCCATCAACCGTTGAGCA	7680
DB	7621	GGGCTCTGATCAGGCATGCGCTGCGGAGGAAAGTAAAGCTGCCCATCAACCGTTGAGCA	7680
QY	7681	ACTCTTGTGCGTCAACCAACATGCTTACGCCACAACATCCCGCAGCGCAAGCCTCC	7740
DB	7681	ACTCTTGTGCGTCAACCAACATGCTTACGCCACAACATCCCGCAGCGCAAGCCTCC	7740
QY	7741	GSCAGAAAGTCACTTTGACAGATTGCAAGTCTCTGATCATATTACCAGGAGCTAC	7800
DB	7741	GSCAGAAAGTCACTTTGACAGATTGCAAGTCTCTGATCATATTACCAGGAGCTAC	7800
QY	7801	TCAGGAGATGAAGCGGAGGCGTCCACAGTTTAAAGCTTAAAGCTTCTATCTATAGAGGAG	7860
DB	7801	TCAGGAGATGAAGCGGAGGCGTCCACAGTTTAAAGCTTAAAGCTTCTATCTATAGAGGAG	7860
QY	7861	CCTGCAAGCTACGCCCCACATTCGGGCAAAATCCAAATTTGGCTATGGGCAAAAGGACG	7920
DB	7861	CCTGCAAGCTACGCCCCACATTCGGGCAAAATCCAAATTTGGCTATGGGCAAAAGGACG	7920
QY	7921	TCCGGAACCTATCCAGAGGCGGCTTACCACTCCGCTCCGCTGGGAGGACTTCTCTGG	7980
DB	7921	TCCGGAACCTATCCAGAGGCGGCTTACCACTCCGCTCCGCTGGGAGGACTTCTCTGG	7980
QY	7981	AAGACACTGAAACACCAATTTGACACCCATCATGGCAAAAAGTGAGGTTTCTTCTGCTCC	8040

Db 7981 |||||AAGACACTGAAACACCAATTTGACACCAATATGCGAAAGAGTTTCTGCGTCC 8040
Qy 8041 AACAGAGAGGAGGCGCGCAAGCCAGCTCGCCTTATCGTATTCCAGACACTGGGAGTTC 8100
Db 8041 AACAGAGAGGAGGCGCGCAAGCCAGCTCGCCTTATCGTATTCCAGACACTGGGAGTTC 8100
Qy 8101 GTGTATGCGAGAGATGCGCCCTTTACGACGTGTCTCCACCCCTTCTCAGGCCGTGATGG 8160
Db 8101 GTGTATGCGAGAGATGCGCCCTTTACGACGTGTCTCCACCCCTTCTCAGGCCGTGATGG 8160
Qy 8161 GCTCCTCATACGATTTCAATACCTCCCCCAAGCAGCGGGTCGAGTTCTCCTGCTGAATACCT 8220
Db 8161 GCTCCTCATACGATTTCAATACCTCCCCCAAGCAGCGGGTCGAGTTCTCCTGCTGAATACCT 8220
Qy 8221 GGAATCAAAAGAAATGCGCTTATGGGCTTCTCATATGACACCCGCTGTTTTGACTCAACGG 8280
Db 8221 GGAATCAAAAGAAATGCGCTTATGGGCTTCTCATATGACACCCGCTGTTTTGACTCAACGG 8280
Qy 8281 TCACTAGAGTGACATTCGTGTTGAGAGTCAATTTACCAATGTTGTAATTGCGCCCGG 8340
Db 8281 TCACTAGAGTGACATTCGTGTTGAGAGTCAATTTACCAATGTTGTAATTGCGCCCGG 8340
Qy 8341 AGCCACAGAGGCATTAAGTTCGTACAGAGCGGCTTTACATCGGGGTCCTCGACTA 8400
Db 8341 AGCCACAGAGGCATTAAGTTCGTACAGAGCGGCTTTACATCGGGGTCCTCGACTA 8400
Qy 8401 ACTCAAAAGGCGAAGCTCGGCTTATCGCGGTGCGCGCAAGTGGCGTGTACGACTA 8460
Db 8401 ACTCAAAAGGCGAAGCTCGGCTTATCGCGGTGCGCGCAAGTGGCGTGTACGACTA 8460
Qy 8461 GCTGCGGTAAATACCTCACAATGTTACTGAAAGGCACTGAGGCTGTCGAAAGC 8520
Db 8461 GCTGCGGTAAATACCTCACAATGTTACTGAAAGGCACTGAGGCTGTCGAAAGC 8520
Qy 8521 TCCAGGACTGACAGTCTGTTGAACGAGACGACCTTGTGTTATCTGTGAAAGCGCG 8580
Db 8521 TCCAGGACTGACAGTCTGTTGAACGAGACGACCTTGTGTTATCTGTGAAAGCGCG 8580
Qy 8581 GAACCCAGGAGGATGCGCGGCGCTTACGAGGCTTACGAGGCTATGACTAGGTATTCCG 8640
Db 8581 GAACCCAGGAGGATGCGCGGCGCTTACGAGGCTTACGAGGCTATGACTAGGTATTCCG 8640
Qy 8641 CCCCCCGGGGATTCGCCCAACAGAAATACGACCTGAGCTGATAACATCATGTTCTT 8700
Db 8641 CCCCCCGGGGATTCGCCCAACAGAAATACGACCTGAGCTGATAACATCATGTTCTT 8700
Qy 8701 CCAATGTGTCGTGCGGACGATGCTGCGCAAAAGGATATCTACTCACCCTGAC 8760
Db 8701 CCAATGTGTCGTGCGGACGATGCTGCGCAAAAGGATATCTACTCACCCTGAC 8760
Qy 8761 CCACCACCCCTTGCAGGCGTGGTGGGAGACAGCTAGACACACTCCAATCAACTCTT 8820
Db 8761 CCACCACCCCTTGCAGGCGTGGTGGGAGACAGCTAGACACACTCCAATCAACTCTT 8820
Qy 8821 GGCTAGGCAATATCATATGATGCGCCACCTTATGGGCAAGGATGATTCGATGACTC 8880
Db 8821 GGCTAGGCAATATCATATGATGCGCCACCTTATGGGCAAGGATGATTCGATGACTC 8880
Qy 8881 ACTTTTCTCCATCTCTTACGCTCAAGAGCAACTTTGAAAGGCCCTGGAATGTCAGATCT 8940
Db 8881 ACTTTTCTCCATCTCTTACGCTCAAGAGCAACTTTGAAAGGCCCTGGAATGTCAGATCT 8940
Qy 8941 ACGGGCTTGCTACTCCATTGAGCCACTTACCTACCTCAGATCATTTGAACGCTCCATG 9000
Db 8941 ACGGGCTTGCTACTCCATTGAGCCACTTACCTACCTCAGATCATTTGAACGCTCCATG 9000
Qy 9001 GTCTTAGGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGCTGGCTTCAT 9060
Db 9001 GTCTTAGGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGCTGGCTTCAT 9060
Qy 9061 GCCTCAGGAAACTTTGGGTACCACTTCGTAAGCTGAGACATCGGGCCAGAGTGTCC 9120

Db 9061 GCCTCAGGAAACTTTGGGTACCACTTGGAACTTGGAGACATCGGGCCAGAGTGTCC 9120
Qy 9121 GGCCTAAGCTACTGTCCAGGGGGAGGGCGCCCACTTGTGCGACATACCTCTTTAACT 9180
Db 9121 GGCCTAAGCTACTGTCCAGGGGGAGGGCGCCCACTTGTGCGACATACCTCTTTAACT 9180
Qy 9181 GGCAGTAAGGACCAAGCTTTAAACTCACTCCAAATCCGGCGCGTCCAGCTGGACTTGT 9240
Db 9181 GGCAGTAAGGACCAAGCTTTAAACTCACTCCAAATCCGGCGCGTCCAGCTGGACTTGT 9240
Qy 9241 CTGGCTGGTTCTGTCGTGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTGCCC 9300
Db 9241 CTGGCTGGTTCTGTCGTGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTGCCC 9300
Qy 9301 GACCCGCTGGTTCTCGTTGTCCTACTCTCTACTTTCTGTAGGGTAGGCAATTTACCTGC 9360
Db 9301 GACCCGCTGGTTCTCGTTGTCCTACTCTCTACTTTCTGTAGGGTAGGCAATTTACCTGC 9360
Qy 9361 TCCCCAACCGATGAACGGGAGCTAACCACTCCAGGCGCTTAAGCCATTTCTGTTTTT 9420
Db 9361 TCCCCAACCGATGAACGGGAGCTAACCACTCCAGGCGCTTAAGCCATTTCTGTTTTT 9420
Qy 9421 TT 9480
Db 9421 TT 9480
Qy 9481 TTTTTCCTTTCTTTAATGTTGCTCCATCTTAGCCCTAGTCACGGCTAGCTGTGAAAGGT 9540
Db 9481 TTTTTCCTTTCTTTAATGTTGCTCCATCTTAGCCCTAGTCACGGCTAGCTGTGAAAGGT 9540
Qy 9541 CCCTGAGCGCATGACTGCGAGAGTCTGATACCTGCGCTCTCTGCAGATCATGT 9595
Db 9541 CCCTGAGCGCATGACTGCGAGAGTCTGATACCTGCGCTCTCTGCAGATCATGT 9595

RESULT 6
ADR82189
ID ADR82189 standard; DNA; 9587 BP.
XX
AC ADR82189;
XX
DT 16-DEC-2004 (first entry)
XX
DE Hepatitis C virus type 1b polyprotein DNA.
XX
KW antilipemic; cardiant; vasotropic; antiarteriosclerotic; antidiabetic;
KW cytotatic; anticonvulsant; nootropic; muscula; anti-HIV;
KW RNA interference; iRNA; antisense technology; lipid metabolism;
KW cholesterol imbalance; dyslipidaemia hypercholesterolaemia;
KW coronary artery disease; CAD; coronary heart disease; CHD;
KW atherosclerosis; hepatic glucose production;
KW glucose-metabolism-related disorder; diabetes; cancer; breast cancer;
KW colon cancer; lung cancer; neurological disease; Huntington disease;
KW spinocerebellar ataxia; viral disease; AIDS; HCV; polyprotein; gene; ds.
XX
Hepatitis C virus.
OS
XX
PN WO2004080406-A2.
XX
PD 23-SEP-2004.
XX
PP 08-MAR-2004; 2004WO-US007070.
XX
PR 07-MAR-2003; 2003US-0452682P.
PR 12-MAR-2003; 2003US-0454265P.
PR 13-MAR-2003; 2003US-0454982P.
PR 13-MAR-2003; 2003US-0455050P.
PR 14-APR-2003; 2003US-0462894P.
PR 17-APR-2003; 2003US-0463772P.
PR 25-APR-2003; 2003US-0465665P.
PR 25-APR-2003; 2003US-045802P.
PR 09-MAY-2003; 2003US-0459612P.
PR 08-AUG-2003; 2003US-0493986P.

PR	11-AUG-2003; 2003US-0494597P.
PR	26-SEP-2003; 2003US-0506341P.
PR	09-OCT-2003; 2003US-0510246P.
PR	10-OCT-2003; 2003US-0510318P.
PR	07-NOV-2003; 2003US-0518453P.
XX	(ALNY-) ALNYLAM PHARM.
PA	
PI	Manoharan M, Bumcrot D;
PX	WPI; 2004-677362/66.
DR	
DX	Interference RNA agent useful for treating dyslipidemias, coronary artery disease, diabetes, cancer or neurological disease, comprises sense sequence and antisense sequence which has specific modifications.
PT	
PS	Example 5; SEQ ID NO 6688; 378pp; English.
XX	
CC	The invention describes a RNA interference (RNA) agent (I) comprising a sense sequence and an antisense sequence, where the sense sequences have one or more asymmetrical 2'-O alkyl modifications, the antisense sequences have one or more asymmetrical phosphorothioate modifications and the antisense sequence targets a human gene sequence. Also described are: a pharmaceutical preparation comprising (I); reducing (MI) apob-100 levels or glucose-6-phosphatase levels in a subject; producing (I); stabilising (I), involves selecting a sequence with activity and introducing one or more asymmetrical modification in the sequence, where the modification decreases nuclease sensitivity while not decreasing its activity; a kit comprising (I) and instruction for its use; and a device that can be dispense or administer a composition comprising (I). (I) is useful for reducing apob-100 levels or glucose-6-phosphatase levels. (MI) is useful for reducing apob-100 levels or glucose-6-phosphatase levels. (MI) The subject is suffering from a disorder characterised by elevated or otherwise unwanted expression of apob-100, elevated or otherwise unwanted levels of cholesterol, and/or dysregulation of lipid metabolism. The disorder is chosen from the HDL/LDL cholesterol imbalance, dyslipidaemias, hypercholesterolaemia, statin-resistant hypercholesterolaemia, coronary artery disease (CAD), coronary heart disease (CHD) and atherosclerosis. (I) is administered to a subject to inhibit hepatic glucose production or for treating glucose-metabolism-related disorder e.g. diabetes or type-2 diabetes. (I) is useful for treating the diseases as mentioned above, cancer (e.g. breast, colon or lung cancer), neurological disease (e.g., Huntington disease or spinocerebellar ataxia) or viral disease (e.g., AIDS). This sequence represents hepatitis C virus type 1b polyprotein DNA.
XX	
SQ	Sequence 9587 BP; 1921 A; 2865 C; 2714 G; 2087 T; 0 U; 0 Other;
Query Match 87.0%; Score 8348.2; DB 13; Length 9587;	
Best Local Similarity 92.0%; Pred. No. 0;	
Matches 8824; Conservative 0; Mismatches 763; Indels 8; Gaps 1	
Qy	1 GCACGCCCCCTGATGGGGCGACACTCACCATGAATCACTCCCCTGTGAGGAACCTACTG 60
Db	1 GCACGCCCCCGATTGGGGCGCACACTCCACCATAGATCACTCCCCTGTGAGGAACCTACTG 60
Qy	61 TCCTTCACGCAGAAACGGTCTACGCCATGCGCTTAGTGATGTCGTGCACGCTCCAGGAC 120
Db	61 TCCTTCACGCAGAAACGGTCTAGCCATGCGCTTAGTGATGTCGTGCACGCTCCAGGAC 120
Qy	121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGCGAACCGGTGAGTACACCGGAATTGCCAG 180
Db	121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGHAACCGGTGAGTACACCGGAATTGCCAG 180
Qy	181 GACGACCGGGTCTTTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGGCGTGCCTCC 240
Db	181 GACGACCGGGTCTTTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGGCGTGCCTCC 240
Qy	241 GCGAGACTCTAGCCGAGTAGTGTGGGTCCGAAAGCGCCTTGTGGTACTGCTGTATAGS 300
Db	241 GCGAGACTCTAGCCGAGTAGTGTGGGTCCGAAAGCGCCTTGTGGTACTGCTGTATAGS 300
Qy	301 GTGCTTTGCGAGTGCCTCCGGGAGGCTCTCGTAGACCGGTGCACCATGAGCAGCAATCTCTTAAC 360

Db	1381	TGTTGGCGGGGGCCCCACTGGGGTGTCTCTAGCGGGCTTGCCTACTATTCCATTGTTGGGGA	1440
Qy	1441	ACTGGGCTAAGGTTCTGATTTGTGGCGCTACTCTTTTCCCGCGCGTTGACGGGGAGAACCCACA	1500
Db	1441	ACTGGGCTAAGGTTCTGATTTGTGATGCTACTCTTTTGTGGCGTTGACGGGCAACCCACG	1500
Qy	1501	CGACGGGGAGGGTGGCGGGCCACACACTCCGGGTTACGTTCCCTTTTCTCATCTGGGG	1560
Db	1501	TGACAGGGGGAAGGGTAGCCTCCAGACCCAGAGCCTCGTGTCTTGGCTCTCTACAAGGGC	1560
Qy	1561	CGTCTCAGAAAAATCCAGCTTGTCAATACCAACGGCAGCTGGGCACATCAACAGAGCTGCCCC	1620
Db	1561	CATCTCAGAAAAATCCAACTGTGAAACCAACGGCAGCTGGGCACATCAACAGGACCGCTC	1620
Qy	1621	TAAATTGCAATGACTCCCTCCAAACTGGGTTCTTTCGCGCTGTGTTTTTACGCACACAAGT	1680
Db	1621	TGAATTGCAATGACTCCCTCCAAACTGGGTTCAATTTGCTGGCTGTCTACGCACACAGGT	1680
Qy	1681	TCAACTCGTCCGGGTGCCCGAGGCGCATGSCCAGCTGCCGCCCCCATTTGACTGGTTCGCCCC	1740
Db	1681	TCAACGCGTCCGGATGTCCAGAGCGCATGCCCAGCTGCCGCCCCCATTCGACAAAGTTTCGCTC	1740
Qy	1741	AGGGTGGGGCCCCATCACCTATACTAAGCCTTAACAGCTCGGATCAGAGGCCCTTATTGCT	1800
Db	1741	AGGGTGGGGTCCCATCACTCACTGTTGTGCTTAACATCTCGACACAGAGCCCTTATTGCT	1800
Qy	1801	GGCAATACGGGCTCCGACCGGTGTGTGTCGTACCCGCGTCGCAGGTGTGTGTGTCACAGTGT	1860
Db	1801	GGCAATACGCCCAACCGTGGGTATTGTACCCGCGTCGCAGGTGTGTGGCCCAAGTGT	1860
Qy	1861	ATTGTTTACCCCAAGCCCTTGTGTGTGGGACACCGGATCGTTCCGGTGTCCCTACAGT	1920
Db	1861	ATTGCTTACCCCGAGTCTGTGTGTGGTGGGACGACCGACCGTTCCGGAGTCCCAAGT	1920
Qy	1921	ATAGCTGGGGGAGAAATGACACAGAGCTGATGCTCCTCAACACACGCGTCGSCACAAG	1980
Db	1921	ATAGCTGGGGGAGAAATGACACAGAGCTGCTGTACTCAACACACGCGGCCGCCGCAAG	1980
Qy	1981	GCAACTGTTCCGGCTGACATGGAATGAATAGTACTGGGTTCACTAAGACGTCGCGAGGTC	2040
Db	1981	GCACTGTTTCGGCTGTACATGGATGAATAGCACCGGTTCAACAGACGTGCGGGGCC	2040
Qy	2041	CCCCGTGTAAACATCGGGGGGTCCGTAAACCGCACCTTGATCTGCCCCACGGACTGCTTCC	2100
Db	2041	CCCCGTGTAAACATCGGGGGGTGGCAACAACACCTTGATTTGCCCCACCGAATGCTTCC	2100
Qy	2101	GGAGCACCCCGAGGCTACTTTACAAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT	2160
Db	2101	GAAGACACCCCGAGGCCACTTACACCAAAATGCGGCTCGGGTCTTGGTTGACACCTAGGT	2160
Qy	2161	GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTCGCACTCTCAATTTTTTTCATCT	2220
Db	2161	GTCTAGTTGACTACCCATACAGACTTTTGGCACTACCCCTGCACTATCAATTTTACCATCT	2220
Qy	2221	TTAAGGTTAGGATGTATGTGGGGGCGTGGAGCACAGGCTCAATGCGCATGCAATTTGGA	2280
Db	2221	TCAAGGTCAGGATGTATGTGGGGGCGTGGAGCACAGGCTCAACGCGCGTGCAATTTGGA	2280
Qy	2281	CTCGAGGAGCGGCTGTAACTTTGAGGACAGGGATAGGTGAGAACTCAGCCCGCTGCTGC	2340
Db	2281	CCCGAGGAGCGGCTGTGACCTTGGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTAT	2340
Qy	2341	TGCTCTAACACAGATGGCAGATACTGCCCTGTGTTTTACACCCCTACCGGCTTTATCCA	2400
Db	2341	TGTCTAACACGGATGGCAGGTACTGCCCTGTTCTTTACACCCCTACCGGCTCTGTCCA	2400
Qy	2401	CTGGTTTGATCCATCTCCATCAGAAACATCGTGGACGTGCAATCCTGTACGGTGTAGGGT	2460
Db	2401	CTGGATTGATCCACCTCCATCAGAATATCTGTGGAGCTGCAATCCTGTACGGTGTAGGGT	2460
Qy	2461	CAGCGTTTGTCTCTTTTGCATCAAAATGGGAGTACATCTCTGTGTGTTTTCTCTCTCTGG	2520
Db	2461	CAGTGGTTGTCTCGGTCGTTAATCAAAATGGGAGTATGTTCTCTGCTCTTCTCTCTCTGG	2520

Db 5761 TCCTGTTTAAATCTTTGGGGGGTGGGTGGCTGCCCAATCGCCCCCGGCGCTT 5820
Qy 5821 CAGCTTTCTGGCGCGGCATCGCGGAGCGGCTGTGTGCAGCATAGGCTTTGGGAAG 5880
Db 5821 CGGCTTTCTGGCGCGGCATCGCGGTGCGGCTGTGTGCAGCATAGGCTTTGGGAAG 5880
Qy 5881 TGTCTGTGACATCTTTGGCGGGCTATGGGCGAGGGTAGCCGCGCACTCGTGGCTTTA 5940
Db 5881 TGTCTGTGACATCTTTGGCGGGTATGGAGCAGAGGTGGCGCGGCTGTGTGGCTTTA 5940
Qy 5941 AGTTCATAGCGCGAGGTGCCCTCCACCGAGACCTGGTCAACTTACTCCCTGCCATCC 6000
Db 5941 AGTTCATAGCGCGAGATGCCCTCTACCGAGACCTGTGTCAATCTACTCTCTGCCATCC 6000
Qy 6001 TCTCTCTGTGTCCTGTGTCGGGTCTGTGCGCAGCAATACTGCGTGGCACTGG 6060
Db 6001 TCTCTCTGTGTCCTGTGTCGGGTCTGTGTCAGCAATACTGCGTGGCACTGG 6060
Qy 6061 GCCGGGAGAGGGGCTGTGCAGTGAATGAACCGGCTGTAGGTTTCGGCTTCGCGGGTA 6120
Db 6061 GTCCGGAGAGGGGCTGTGCAGTGAATGAACCGGCTGTAGGTTTCGGCTTCGCGGGTA 6120
Qy 6121 ACCACGCTCCCTAGCAGTATGTGCTCAGAGCGAGCTGCAGACAGTGTCACTCAGA 6180
Db 6121 ATCAGGTTTCCCCACGCACTATGTGCTGAGAGCGAGCGCGAGCGGTACTCAGA 6180
Qy 6181 TCCTCTCAGCCTTACCATCACTCAACTGTCTGAAGCGGCTCCACAGTGGATTAAGAG 6240
Db 6181 TCCTCTCAGCCTTACCATCACTCAGCTGTGTAAGAGGCTCCACAGTGGATTAAGAG 6240
Qy 6241 ACTGCTCTAGCCATGTCCTGGCTGTGGCTTAGGGATGTTTGGGATGTGATGACGG 6300
Db 6241 ACTGCTCTAGCCATGTCCTGGCTGTGGCTTAGGGATGTTTGGGATGTGATGACGG 6300
Qy 6301 TGTGACTGACTTCAAGACCTGGCTCAGTCCCAACTCTTCGCGCGGTACCGGGAGTCC 6360
Db 6301 TGTGACTGACTTCAAGACCTGGCTCAGTCCCAACTCTTCGCGCGGTACCGGGAGTCC 6360
Qy 6361 CTTTCTCTCATGCCAACGGGGTACAAAGGAGTCTGCGCGGGGACGGCATATGCAAA 6420
Db 6361 CTTTCTCTCATGCCAACGGGGTACAAAGGAGTCTGCGCGGGGACGGCATATGCAAA 6420
Qy 6421 CCACCTCCCATCGGAGCAGATCGCCGAGCATGTCAAAAAGGTTTCAATGAGATCG 6480
Db 6421 CCACCTCCCATCGGAGCAGATCAACGAGCATGTCAAAAAGGTTTCAATGAGATCG 6480
Qy 6481 TAGGCGCTAGAACCTGCAGACACAGTGGGACGGAAGTTCCTCCCATCAACGATACCA 6540
Db 6481 TCGGCGCTAAGACCTGCAGCAACACGTGGCATGGAACATTTCCCATCAACGATACCA 6540
Qy 6541 CGGACCTTGCACACCTCCCGCGCCCAACTATTCAGGGCGCTATGGCGGTGGCTG 6600
Db 6541 CGGCGCTTGCACACCTTCTCAGCGCCAACTATTTAGGGCGCTGTGGCGGTGGCGG 6600
Qy 6601 CTGAGGAGTACGTGGAGTTACGGGTGTGGGGATTTCCACTACGTGACGGGATGACCA 6660
Db 6601 CTGAGGAGTACGTGGAGTTACGGGTGTGGGGATTTCCACTACGTGACGGGATGACCA 6660
Qy 6661 GTGCAACAGTAAAGTGCCCATGCCAGTTCGGGCCCGGCAATTTCTTACGAGGTGGATG 6720
Db 6661 GTGCAACAGTAAAGTGCCCATGCCAGTTCGGGCCCGGCAATTTCTTACGAGGTGGAG 6720
Qy 6721 GAGTGGGTTGCACAGGTACGCTCGGCGTGCMAACCTTCTACGGGAGGACGTACGT 6780
Db 6721 GAGTGGGTTGCACAGGTACGCTCGGCGTGCAGGCGCTCTCTACGGGAGGAGTTACAT 6780
Qy 6781 TCCAGGTCGGGCTCAACCAATACTTGTTCGGGTTCGAGCTCCCATGCGAGCCCAACCG 6840
Db 6781 TCCAGGTCGGGCTCAACCAATACTTGTTCGGGTTCGAGCTCCCATGCGAGCCCAACCG 6840
Qy 6841 ACGTAACAGTCTTACTTCCATGTCTACCGATTCCTCCCATCAATTAACAGAGACGGTA 6900
Db 6841 ATGTAGCAGTGTCTACTTCCATGTCTACCGACCCCTCCCATCAACAGAGACGGTCTA 6900

Qy 6901 ACGTAGGCTGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCAGTTGT 6960
Db 6901 ACGTAGGTTGGCAGGGGCTCTCCCCCTCTTGTGCCAGCTCTTTCAGCTAGCAGTTGT 6960
Qy 6961 CTGCGGCTTTCTTTGAAGGCGACATGACCTACCCACCATGACTCCCGGAGCGCTGACTCA 7020
Db 6961 CTGCGGCTTTCTTTGAAGGCGACATGACCTACCCACCATGACTCTCTCGGAGCTGACTCA 7020
Qy 7021 TCGAGGCCAACCTTTGTGGCGGAGGAGATGGGCGGAAACATCACTCGGTGGAGTCAG 7080
Db 7021 TCGAGGCCAACCTCTGTGGCGGAGGAGATGGGCGGAAACATCACTCGGTGGAGTCAG 7080
Qy 7081 AGNATAGGTAGTAATTTCTGGACTTTTGAACCGCTTCAACCGGAGGGGAGTAGAGGG 7140
Db 7081 AGNATAGGTAGTGTAGTCTTGGACTTTTGAACCGCTTCAACCGGAGGGGAGTAGAGGG 7140
Qy 7141 AGATATCCGTGCGGCGGAGATCTTCGGAATAATCCAGGAAGTTCCCTCAGCGGTGCCA 7200
Db 7141 AGATATCCGTTCGCGGAGATCTTCGGAATAATCCAGGAAGTTCCCTCAGCGGTGCCA 7200
Qy 7201 TATGGGACCGCCGAGTACAACTCTCTCACTGTCTAGAGTCTTGAAAGGACCCGAGTACG 7260
Db 7201 TCTGGCGCGCCGGATTACAACTCTCACTGTCTAGAGTCTTGAAAGGACCCGAGTACG 7260
Qy 7261 TCCTTCGGTGTACAGGATGCCATTCACCTACCAAGGCTCTTCCCAATACCACCTC 7320
Db 7261 TCCTTCGGTGTGACAGGATGCCATTCACCTACCAAGGCTCTTCCCAATACCACCTC 7320
Qy 7321 CACGGAAGAGAGAGGTTGTCTCTGACAGAAATCCAAATGTCTTCTGCTTGGCGGAGC 7380
Db 7321 CACGGAAGAGAGAGGTTGTCTCTGACAGAGTCTTCAACAGAGTCTTCTGCTTACGCGAGC 7380
Qy 7381 TCGCCACTAAGACCTTCGGTGTCTGCGGATGTCTGCGCGGTGTAGAGCGGACGCGACCG 7440
Db 7381 TCGCCACTAAGACCTTCGGTGTCTGCGGATGTCTGCGCGGTGTAGAGCGGACGCGACCG 7440
Qy 7441 CCCTTCTGACTTGGCTCCGACGAGCGGTGACAAAGATCCGAGTTGAGTTCGTACTCCT 7500
Db 7441 CCCTTCTGACTTGGCTCCGACGAGCGGTGACAAAGATCCGAGTTGAGTTCGTACTCCT 7500
Qy 7501 CCATGCCCTTTGAAGGGGAGCGCGGAGACCCCGATCTCAGCGAGCGGTCTTGGTCTA 7560
Db 7501 CCATGCCCTTTGAAGGGGAGCGCGGAGACCCCGATCTCAGTGACGGGTCTTGGTCTA 7560
Qy 7561 CCGTGTAGTGGAGGCTAGTGGAGTGTCTGCTCTCAATGTCTCTATACGTGGACAG 7620
Db 7561 CCGTGTAGTGGAGGCTAGTGGAGTGTCTGCTCTCAATGTCTCTACATGGACAG 7620
Qy 7621 GCGCCTGTATCAGCCATGCGTGTGGAGGAAAGTAAAGTGTGCCATCAACCCGTTGAGCA 7680
Db 7621 GCGCCTGTATCAGCCATGCGTGTGGAGGAAAGTAAAGTGTGCCATCAACCCGTTGAGCA 7680
Qy 7681 ACTTTTGTCTGCTCAGCAACATGCTCTACGCCCAACATCCCGAGCGCAAGCTCC 7740
Db 7681 ACTTTTGTCTGCTCAGCAACATGCTCTACGCCCAACATCTCGGAGCGAGCTTCG 7740
Qy 7741 GGCAGAAAGAGTCACTTTTGAAGTGTGCAAGTCTCTGGATGATCATTTACCGGACGTCAC 7800
Db 7741 GGCAGAAAGAGTCACTTTTGAAGTGTGCAAGTCTCTGGATGATCATTTACCGGACGTCAC 7800
Qy 7801 TCAAGGAGATGAAGGCGGAGGCTGCAAGTAAAGGCTTAAGCTTCTATCTATAGAGGAG 7860
Db 7801 TCAAGGAGATGAAGGCGGAGGCTGCAAGTAAAGGCTTAAGCTTCTATCTATAGAGGAG 7860
Qy 7861 CCGCAAGCTGACCGGCGGAGGCTTTCGCGCAATCCAAATTTGGCTATGGGCAAGGAGC 7920
Db 7861 CCGCAAGCTGACCGGCGGAGGCTTTCGCGCAATCCAAATTTGGCTATGGGCAAGGAGC 7920
Qy 7921 TCCGGAACCTATCAGAGGCGGCTTAAACACATCCGCTCCGCTGTGGGAGGACTTGTGG 7980
Db 7921 TCCGGAACCTATCAGAGGCGGCTTAAACACATCCGCTCCGCTGTGGGAGGACTTGTGG 7980

Qy	7981	AAGACACTGAAA	CHC	CAATTTGA	CAC	CAACCAATCAT	TGCGA	AAAAAGTAGAGTTTCTCGGTCC	8040
Db	7981	AAGACACTGTGACACCA	CAATTTGAC	CAACCAATCAT	TGCGA	AAAAAGTAGAGTTTCTGTGTCC	8040		
Qy	8041	AACCAGAGAAGGAGGCG	CAAGCCAGCTCG	CGCTTATCGTATTTCC	CAGACCTCGGAGTTC				
Db	8041	AACCAGAGAAGGAGGCG	CAATTTGAC	CAACCAATCAT	TGCGA	AAAAAGTAGAGTTTCTCGGTGAATACCT	8100		
Qy	8101	GTGTATCGGAGAAGATGG	CCCTTTACGACGTGGTCT	CCACCTTCTC	CAGGCGGTGATGG				
Db	8101	GTGTATCGGAGAAGATGG	CCCTCTATGATGTGGTCT	CCACCTTCTC	CAGGCGGTGATGG				
Qy	8161	GCTCCTCATACGGATTTCA	ATPACTCC	CCCAAGCAGCGGGT	CGAGTTCCTGGTGAATACCT				
Db	8161	GCTCCTCATACGGATTTCA	GTACTCTCTGGG	CAGCGAGTTCGAGTTCCTGGTGAATACCT					
Qy	8221	GGAAATCAAGAANAATG	CCCTATATGGGCTTCT	CATATGACACCCCGT	TTTGACTCAACGG				
Db	8221	GGAAATCAAGAANAATG	CCCTATATGGGCTTCT	CATATGACACCTCGCTGT	TTGACTCAACGG				
Qy	8281	TCACGTAGAGTGACATTT	CGTGTGTGAGGAGTCA	ATTTACCAATGTGTG	ACTTCGGCCCCCG				
Db	8281	TCACGTAGAGTGACATTT	CGTGTGTGAGGAGTCA	ATTTACCAATGTGTG	ACTTCGGCCCCCG				
Qy	8341	AGGCCAGACAGGCCAAT	AAGTTCGTCTC	CAGAGGGCTTTACATCGGGGGT	CCCTCGACTA				
Db	8341	AAGCCAGACAGGCCAAT	AAATTCGTCTC	CAGAGGGCTTTATATCGGGGGT	CCCTCGACTA				
Qy	8401	ACTCAAAAGGGCAGAACT	GCGGTATTCGCGGTG	CCCGCAAGTGGCGTGCTC	GACGACTA				
Db	8401	ATTCAAAAGGGCAGAACT	GCGGTATTCGCGGTG	CCCGCAAGTGGCGTGCTC	GACGACTA				
Qy	8461	GCTCGGTATACCCCTC	ACATGTTTAC	TGTAAGGCCACTG	CAGCGCTGTCGAGCTG	CGAAGC			
Db	8461	GCTCGGTATACCCCTC	ACATGTTTAC	TGTAAGGGCTCTG	CAGCGCTGTCGAGCTG	CGAAGC			
Qy	8521	TCCAGGACTGCACGATG	CTCGTGAACGGACAG	CACTTGTCTGTATCTGTGA	AAAGCGCGG				
Db	8521	TCCAGGACTGCACGATG	CTCGTGAACGGACAG	CACTTGTCTGTATCTGTGA	AAAGCGCGG				
Qy	8581	GAA	CCCCAGGAGATTCGG	CGGCCCTACGAGCGCTT	CACGGAGGCTATGACTAGGTATTCG				
Db	8581	GAA	CCCCAAGAGGACCG	CGCGAGCTTACGAGTCTTACGCGAGGCTATGACTAGGTATTCG					
Qy	8641	CCCCCGGGGATTCG	CCCCCAACAGAAATAC	GNACCTGGAGCTGATAA	CAATCATGTCTCT				
Db	8641	CCCCCGGGGATTCG	CCCCCAACAGAAATAC	GNACCTGGAGCTGATAA	CAATCATGTCTCT				
Qy	8701	CCAATGTGTAGTTCG	CGCGCAGATGCATCT	TGGCAAAAGGGTATAC	TACTCTCACCCGCTGACC				
Db	8701	CCAATGTGTAGTTCG	CGCGCAGATGCATCT	TGGCAAAAGGGTATAC	TACTCTCACCCGCTGACC				
Qy	8761	CCACCA	CCCCCTTTGAC	CGGGCTCGTGGGAGACAG	CTAGACACATCTCCAATCAACTCTT				
Db	8761	CCACCA	CCCCCTTTGAC	CGGGCTCGTGGGAGACAG	CTAGACACATCTCCAATCAACTCTT				
Qy	8821	GGCTAGGCAATATCAT	GATGTATGGGCCCA	CCCTATGGGCAAGAGTATTTCTG	TGACTGACTC				
Db	8821	GGCTAGGCAATATCAT	GATGTATGGGCCCA	CCCTATGGGCAAGAGTATTTCTG	TGACTGACTC				
Qy	8881	ACTTTTCTCCATCTCT	AGCTCAAGAGCAACT	TTGAAAAAGCCCTGGATTGT	GCAGATCT				
Db	8881	ACTTTTCTCCATCTCT	AGCTCAAGAGCAACT	TTGAAAAAGCCCTGGATTGT	GCAGATCT				
Qy	8941	ACGGGGCTTGCTACT	TCCATTGAGCCACT	TGACCTCAGATCATTTGAA	CGACTCCATG				
Db	8941	ACGGGGCTTGCTACT	TCCATTGAGCCACT	TGACCTCAGATCATTTGAA	CGACTCCATG				
Qy	9001	GTCTTAGGCATTTAC	ACTCACAAGTTACT	TCTCCAGGTGAGATCAAT	ATAGGTGGCTTCAT				
Db	9001	GCCTTAGGCATTTTCA	CTCCTCATAGTTACT	TCTCCAGGTGAGATCAAT	ATAGGTGGCTTCAT				
Qy	9061	GCCTCAGAAA	ACTTTGGGGTAC	CA	CCCTTTGCGAACTCTGGAGACAT	TCGGGGCCAGAGTGTCC			

[illegible]

26-SEP-2003; 2003US-0506341P.
09-OCT-2003; 2003US-0510246P.
10-OCT-2003; 2003US-0510318P.
07-NOV-2003; 2003US-0518453P.
08-MAR-2004; 2004WO-US007070.
05-APR-2004; 2004WO-US010586.
(ALNY-) ALNYLAM PHARM INC.
Manoharan M, Elbashir S, Harborth J;
WPI; 2004-766693/75.
New interference RNA agent comprising sense sequence and antisense
PT sequence having cholesterol moieties, useful for reducing apob-100 levels
PT or glucose-6-phosphatase levels.
XX
PS Example 4; SEQ ID NO 6688; 324pp; English.
XX
CC The invention describes an interference RNA (iRNA) agent (I) comprising a
CC sense sequence and an antisense sequence, where the sense sequence
CC comprises one or more cholesterol moieties, and the antisense sequence
CC targets a human gene sequence. The following are disclosed: a
CC pharmaceutical composition comprising (I); and a device for administering
CC (I) into a patient. (I) is useful for reducing apob-100 levels or glucose
CC -6-phosphatase levels in a subject. (I) targets a sequence identical to
CC any one of sequences as given in the specification. (I) comprises a
CC cholesterol moiety. The cholesterol moiety is coupled to a sense strand.
CC (I) further comprises a second cholesterol moiety. The second cholesterol
CC moiety is coupled to a sense strand. (I) has 21 or more nucleotides. The
CC duplex region of (I) is 19 nucleotides in length. The subject is
CC suffering from a disorder having elevated or otherwise unwanted
CC expression of apo-B-100, elevated or otherwise unwanted levels of
CC cholesterol, and/or dysregulation of lipid metabolism. The disorder is
CC chosen from HDL/LDL cholesterol imbalance, dyslipidaemia (e.g., familial
CC combined hyperlipidaemia or acquired hyperlipidaemia), coronary
CC hypercholesterolaemia, statin-resistant hypercholesterolaemia, coronary
CC artery disease, coronary heart disease and atherosclerosis, preferably
CC statin-resistant hypercholesterolaemia. (I) is administered to a subject
CC to inhibit hepatic glucose production or for treating glucose-metabolism-
CC related disorders e.g., type-2 diabetes or glitazone-resistant diabetes.
CC (I) has endonuclease or exonuclease resistance. This sequence represents
CC a polynucleotide associated with liver disorders, expression of which can
CC be altered using interfering RNA's described in the invention.
XX
SQ Sequence 9587 BP; 1921 A; 2865 C; 2714 G; 2087 T; 0 U; 0 Other;
Query Match 87.0%; Score 8348.2; DB 13; Length 9587;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 8824; Conservative 0; Mismatches 763; Indels 8; Gaps 1;
Qy 1 GCCAGCCCCGTGATGGGGGCGACACTCCACCATGAATCACCCTCTGTGAGGAATCTACTG 60
Db 1 GCCAGCCCCGATTTGGGGGCGACACTCCACCATAGATCACCCTCTGTGAGGAATCTACTG 60
Qy 61 TCTTACGACGAGAGCGTCTAGCCATGCGCTTAGTATGAGTCTCGTCAGCGCTCCAGGAC 120
Db 61 TCTTACGACGAGAGCGTCTAGCCATGCGCTTAGTATGAGTCTCGTCAGCGCTCCAGGAC 120
Qy 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATTGCCAG 180
Db 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATTGCCAG 180
Qy 181 GACGACCGGGTCTTTCTTGGATCAACCGCTCAATGCTGAGATTTGGGGTGGCCCCC 240
Db 181 GACGACCGGGTCTTTCTTGGATCAACCGCTCAATGCTGAGATTTGGGGTGGCCCCC 240
Qy 241 GCGAGACTGCTAGCGAGTAGTGTGGGTGCGGAAAGGCGCTTGTGGTACTGCTGATAGG 300
Db 241 GCGAGACTGCTAGCGAGTAGTGTGGGTGCGGAAAGGCGCTTGTGGTACTGCTGATAGG 300
Qy 301 GTGCTTGCAGTGCCTCCCGGAGGCTCTCGTAGACCGTGCACCATGACGCAATCTCAAC 360
Db 301 GTGCTTGCAGTGCCTCCCGGAGGCTCTCGTAGACCGTGCATCATGAGCACAATCTCAAC 360
Qy 361 CTCACAGAAAAACAAACGTAACACAAACCGCCGCCACAGGACGTAAGTTCCTCCGGCG 420
Db 361 CTCACAGAAAAACAAACGTAACACAAACCGCCGCCACAGGACGTAAGTTCCTCCGGCG 420
Qy 421 GTGGTCAGATCGTGTGGTGGAGTTTACCTGTGTGCGGCGCAGGGGCCAGGTGGGTGTGC 480
Db 421 GTGGTCAGATCGTGTGGTGGAGTTTACCTGTGTGCGGCGCAGGGGCCAGGTGGGTGTGC 480
Qy 481 GCGCGCTAGGAAAGGCTTCGAGCGGTGCAACCTCGTGGAAAGGCGCAACCTATCCCAA 540
Db 481 GCGCGCTAGGAAAGGCTTCGAGCGGTGCAACCTCGTGGAAAGGCGCAACCTATCCCAA 540
Qy 541 AGGCTGCGGACCCGAGGCGAGGCTGGGCTCAGCCCGGGTACCTTGGCCCCCTCTATG 600
Db 541 AGGCTGCGGACCCGAGGCGAGGCTGGGCTCAGCCCGGGTACCTTGGCCCCCTCTATG 600
Qy 601 GCAATGAGGCGCTGGGCTGGGCGAGGATGGCTCTGTCAACCCCGGGCTCCCGGCTAGTT 660
Db 601 GCAATGAGGCGCTGGGCTGGGCGAGGATGGCTCTGTCAACCCCGGGCTCCCGGCTAGTT 660
Qy 661 GGGGCCCCACGAGACCCCGGCGTAGGTTCGGCTAACTTTGGGTAAAGTTCATGATACCTTA 720
Db 661 GGGGCCCCACGAGACCCCGGCGTAGGTTCGGCTAACTTTGGGTAAAGTTCATGATACCTTA 720
Qy 721 CATGCGGCTTCGCGGATCTCATGGGTGACATTCGGCTCGTGGGCGCCCCCTAGGGGCG 780
Db 721 CATGCGGCTTCGCGGATCTCATGGGTGACATTCGGCTCGTGGGCGCCCCCTAGGGGCG 780
Qy 781 CTGCGAGGCGCTTGGGACACGCTGTCGGGTTCCTGGAGGACGGGTGAATATGCAACAG 840
Db 781 CTGCGAGGCGCTTGGGACACGCTGTCGGGTTCCTGGAGGACGGGTGAATATGCAACAG 840
Qy 841 GGAATCTGCGCGGTGCTCTTTCTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 841 GGAATCTGCGCGGTGCTCTTTCTCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Qy 901 TCCAGCTTCGCTTATGAGTGGCGAAGTGTCCGGGATATACATGTACGAGACGACT 960
Db 901 TCCAGCTTCGCTTATGAGTGGCGAAGTGTCCGGGATATACATGTACGAGACGACT 960
Qy 961 GCTTCAACTCAGCATTTGTATGAGGCGGACATGATCATCATCTATCTCCCGGTGCG 1020
Db 961 GCTTCAACTCAGCATTTGTATGAGGCGGACATGATCATCATCTATCTCCCGGTGCG 1020
Qy 1021 TGCCCTGTGTTACGAGGCTAAACAGCTCCCGTCTGGGTAGCGCTCACTCCACGCTCG 1080
Db 1021 TGCCCTGTGTTACGAGGCTAAATTTCTCCCGTCTGGGTAGCGCTCACTCCACGCTCG 1080
Qy 1081 CGGCGAGGAATGCGAGCTCCCACTACGAAATACGACGCGCAGCTGCTGCTGTTG 1140
Db 1081 CGGCGAGGAATGCGAGCTCCCACTACGAAATACGACGCGCAGCTGCTGCTGTTG 1140
Qy 1141 GGAGCGGTGCTTCTGCTCCGCTATGACGTGGGGGATCTCTGGGATCTATTTCTCTCG 1200
Db 1141 GGAGCGGTGCTTCTGCTCCGCTATGACGTGGGGGATCTCTGGGATCTATTTCTCTCG 1200
Qy 1201 TCTCCAGCTGTTTCACTTCTCGCTCGCGCATGACAGTGCAGGACTGCAACTGCT 1260
Db 1201 TCTCCAGCTGTTTCACTTCTCACTCGCGCATGACAGTGCAGGACTGCAACTGCT 1260
Qy 1261 CAATCTATCCCGGCGCATGATCAGTCAACCGATGGCTTGGGATATGATGAACTGCT 1320
Db 1261 CAATCTATCCCGGCGCATGATCAGTCAACCGATGGCTTGGGATATGATGAACTGCT 1320
Qy 1321 CACTTAAACAGCCCTAGTGTGCGAGTTGCTCCGGATCCCAACAGCTGCTGTTGACA 1380
Db 1321 CACTTAAACAGCCCTAGTGTGCGAGTTGCTCCGGATCCCAACAGCTGCTGTTGACA 1380
Qy 1381 TGGTGGCGGGGCGGCGGAGTCTGGGGGCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
Db 1381 TGGTGGCGGGGCGGCGGAGTCTGGGGGCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCT 1440

QY 1441 ACTGGCTAAGGTTCTGATTGTGGCGCTACTCTTTGGCGGCGTTGACGGGAGACCCACA 1500
DB |||||
QY 1441 ACTGGGCTAAGGCTTGTATTGTGATGCTACTCTTTGCTGGCGTGAACGGGCACACCCAGC 1500
DB |||||
QY 1501 CGACGGGAGGGTGGCCGCCACACACCTCCGGGTTACAGTCCCTTTTCTCATCTGGGG 1560
DB |||||
QY 1501 TGACAGGGGGAAGGGTAGGCTCCAGCACCCAGAGCTCGTGTCTCTGCTCTCACAAAGGCG 1560
DB |||||
QY 1561 CGTCTCAGAAATCCAGCTTGTGAATACCAACGGGAGCTGGCACATCAACAGGACTGCC 1620
DB |||||
QY 1561 CATCTCAGAAAAATCCAACCTCGTGAACACCAACGGGAGCTGGCACATCAACAGAACCGCTC 1620
DB |||||
QY 1621 TAAATTTGCAATCACTCCCTCCAAACTGGGTTCTTTGGCGGCTGTTTACGGCACACAAGT 1680
DB |||||
QY 1621 TGAATTTGCAATGACTCCCTCCAAACTGGGTTCAATTTGCTGGCTGTTTACGGCACACAAGT 1680
DB |||||
QY 1681 TCAACTCGTCCGGGTGCCGGAGCGCATGGCCAGCTGCCGCCCAATGACTGTTTCGGCC 1740
DB |||||
QY 1681 TCAACGGCTCCGATGTCCAGAGCGCATGGCCAGCTGCCGCCCAATCGACAAGTTGCTC 1740
DB |||||
QY 1741 AGGGTGGGGCCCCATCACTATATACTAAGCCTAAACAGCTTCGGATCAGAGGCCCTTATGCT 1800
DB |||||
QY 1741 AGGGTGGGGTCCCATCACTCACTGTTGCTTAAACATCTCGACACAGAGCCCTTATGCT 1800
DB |||||
QY 1801 GGCATTACGGCTCGACCGCTGGTGTGCTACCCGCTCGCAGGTGTGTGTCAGTGT 1860
DB |||||
QY 1801 GGCATTACGACCCCAACCGCTGGTGTGTTACCCGCTCGCAGGTGTGTGTCAGTGT 1860
DB |||||
QY 1861 ATTGTTTACCCCAAGCCCTGTTGTGGTGGGACCAACCGATGCTTCGGGTGCCCTACGT 1920
DB |||||
QY 1861 ATTGCTTACCCCGAGTCTGTTGTGGTGGGAGCAGACCGACCTTCGGGAGTCCCAAGT 1920
DB |||||
QY 1921 ATAGCTGGGGGAGAAATGAGACAGAGTATGCTCTCAACACACAGCGCTCGGCACAAAG 1980
DB |||||
QY 1921 ATAGCTGGGGGAGAAATGAGACAGAGTATGCTCTCAACACACAGCGCGCGCGCAAG 1980
DB |||||
QY 1981 GCNACTGTTCCGCTGTATCATGGATGAATAGTACTGGGTTCACTAAGACGTGGGAGGTC 2040
DB |||||
QY 1981 GCNACTGGTTCGGCTGTATCATGGATGAATAGACACCGGGTTCAACAAAGCTGGGGGGCC 2040
DB |||||
QY 2041 CCCCGTTAAACATCGGGGGGTCGGTAAACCGACCTTGATCTGCCCCACGGACTGCTTCC 2100
DB |||||
QY 2041 CCCCGTTAAACATCGGGGGGTTGGCAACACACCTTGATTTGCCCCACGGATTGCTTCC 2100
DB |||||
QY 2101 GGAAGCACCCCGAGGCTACTTACACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT 2160
DB |||||
QY 2101 GAAAGCACCCCGAGGCACTTACACAAAATGCGGCTCGGGTCTGTTGGTTGACACCTAGGT 2160
DB |||||
QY 2161 GCCTAGTACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTCCATCT 2220
DB |||||
QY 2161 GTCTAGTTGACTACCCATACAGACTTTGGCACTTACCCCTGCACTATCAATTTTACCATCT 2220
DB |||||
QY 2221 TTAAGGTTAGGATGTATGTGGGGGGCTGGAGCACAGGCTCAATGCCGATGCAATTGGA 2280
DB |||||
QY 2221 TCNAGGTCAGGATGTACGTGGGGGGCTGGAGCACAGGCTCAACGCGGCTGCAATTGGA 2280
DB |||||
QY 2281 CTCGAGGAGCGCTGTAACTTCGAGGACAGGATAGGTGAGAACTCAGCCGCTGCTGC 2340
DB |||||
QY 2281 CCCGAGGAGCGCTGTGACCTCGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTAT 2340
DB |||||
QY 2341 TGTCTACACAGAGTGGCAGATGCTGCCCTGTGCTTTTCAACCCCTACCGGCTTTATCCA 2400
DB |||||
QY 2341 TGTCTACACGGAGTGGCAGGTTACTGCCCTGTTCCTTTTACCACTTACCGGCTCTGTCCA 2400
DB |||||
QY 2401 CTGGTTTGATCCATCTCCATCAGAAACATCGTGGACGTGCAATACCTGTACGGGTAGGCT 2460
DB |||||
QY 2401 CTGGATTGATCCACCTTCCATCAGAAATATCGTGAAGCTGCAATACCTGTACGGGTAGGCT 2460
DB |||||
QY 2461 CAGCGTTGTCTCTTGTGCAATCAATGGGAGTACATCTCTGTTGCTTTTCTCTCTCGG 2520
DB |||||
QY 2461 CAGTGGTGTCTCCGTCGTAAATCAAAATGGGAGTATGTTCTGCTGCTCTTCTCTCTCG 2520

QY 2521 CAGACGGCGCGTGTGTGCTCTGTTGTGATGATGCTGCTGATAGCCAGGCTGAGGCGG 2580
DB |||||
QY 2521 CGGACGGCGCGTCTGTGCTGCTGTTGTGATGATGCTGCTGATAGCCAGGCTGAGGCGA 2580
DB |||||
QY 2581 CTTTAGAGAACTTGGTGTCTCAATGCGGCGTCCGTGGCGGAGCGCATGGTATTCTCT 2640
DB |||||
QY 2581 CTTTAGAGAACTTGGTGTCTCAATGCGGCGTCTGTGGCGGAGCGCATGGCCCTTCTCT 2640
DB |||||
QY 2641 CTTTCTTGTGCTTCTTCTGCGCCGCTGCTATCAATTAAGGCGAGGCTGCTCTCTGGGGGG 2700
DB |||||
QY 2641 CTTTCTCTGCTGTTCTTCTGCGCGCTGCTGATCAATCAAGGCGAGGCTGCTCTCTGGGGGG 2700
DB |||||
QY 2701 CGTATGCTTTTTATGGCGTATGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
DB |||||
QY 2701 CATATGCTCTCTATGGCGTATGGCGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
DB |||||
QY 2761 CTTAGCGCTTGGACCGGAGATGGCTGCATCGTGGGGGGTGCGGTCTCTTGTAGTCTGG 2820
DB |||||
QY 2761 CTTATGCCATGGACCGGAGATGGCTGCATCGTGGGAGCGCGGTTTTTGTAGTCTGG 2820
DB |||||
QY 2821 TATTCTTGACCTTGTCACTACTACAAAGTGTCTTCTCCTAGGCTCATATGCTGCTTAC 2880
DB |||||
QY 2821 TACTCTTGACCTTGTCACTACTATTAAGGTGTTCTCTCGTAGGCTCATATGCTGCTTAC 2880
DB |||||
QY 2881 AATACTTTATCACAGAGCCGAGGCGCACATGCAAGTGTGGGTCCCCCTTCAACGTTTC 2940
DB |||||
QY 2881 AATACTTTATCACAGGCGCGAGGCGCACTTGCAAGTGTGGGTCCCCCTCTTAATGTTTC 2940
DB |||||
QY 2941 GGGAGGCGCGGATGCCATCATCTCTCTCAGTGTGGGTTTCATCCAGAGTTAAATTTTG 3000
DB |||||
QY 2941 GGGAGGCGCGGATGCCATCATCTCTCTTACATGCGCGTCCATCCAGAGCTTAATCTTTG 3000
DB |||||
QY 3001 ACATCACAACACTCTGCTCGCATACTCGCCGCTCATGCTGCTCCAGGCTGCATATAA 3060
DB |||||
QY 3001 ACATCACAACACTCTGCTCGCATACTCGGTCCTCATGCTGCTCCAGGCTGCATATAA 3060
DB |||||
QY 3061 CGAGAGTCCGCTACTTCTGTCGCGCTCAAGGCTCATTCGTCATGCTGCTAGTTCGAA 3120
DB |||||
QY 3061 CGAGAGTCCGCTACTTCTGTCGCGCTCAAGGCTCATTCATGCTGCTAGTTCGAA 3120
DB |||||
QY 3121 AAGTCCGCGGGGCTCATTTATGTCMAATGCTTTCATGAAGCTGGGCGGCTGCAGGTA 3180
DB |||||
QY 3121 AGGTCCGCTGGGGCTCATTTATGTCMAATGCTTTCATGAAGCTGGGCGGCTGCAGGTA 3180
DB |||||
QY 3181 CGTACGTTTATAACCATCTTACCCACTCGGGAAGTGGGCCACGCGGCTACGAGACC 3240
DB |||||
QY 3181 CGTACATTTACAACCATCTTACCCCGCTACGGGATTTGGGCCACGCGGCTACGAGACC 3240
DB |||||
QY 3241 TTGCGGTGGCGGTAGAGCCGCTGCTCTTCTCCGCCATGGAGACCAAGGTCATCACCTGGG 3300
DB |||||
QY 3241 TTGCGGTGGCAGTGGAGCCGCTGCTCTTCTCCGATGGAGACCAAGATCATCACCTGGG 3300
DB |||||
QY 3301 GAGCAGACACCGCTCGGTGGGGAATCATCTTGGGCTACCCGCTCTCGGCCGGAAGGG 3360
DB |||||
QY 3301 GAGCAGACACCGCGGCTGTGGGGAATCATCTTGGGCTGCGCCGCTCTCGGCCGGAAGGG 3360
DB |||||
QY 3361 GGAAGGAGATATTTTGGGACCGCTGATGCTCGAAGGCAAGGTCGCGGCTCTCTTG 3420
DB |||||
QY 3361 GAAAGGAGATATCTCTGGGCGCGCGATGATGCTTTGAAGGGCGGGGTTGGGCTCTCTG 3420
DB |||||
QY 3421 CGCCCATACGCGCTACTTCCCAACAAACGCGGGGCTACTTGTGTCATCATCATCTAGCC 3480
DB |||||
QY 3421 CGCCCATACGCGCTACTTCCCAACAGACGCGGGGCTACTTGTGTCATCATCATCTAGCC 3480
DB |||||
QY 3481 TCACAGGCGGGAACAGAACCCAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
DB |||||
QY 3481 TTAACAGGCGGGAACAGAACCCAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
DB |||||
QY 3541 AATCTTCTCGGACCTGTCATCAACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
DB |||||
QY 3541 AATCTTCTCGGACCTGTCATCAACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
DB |||||
QY 3601 CGAAGACCTTAGCCGCTCCAAAGGTTCCNAATCACCCAAATGTACACCAATGTAGACCTGG 3660

Db	3601	CAAAAGACCTTAGCGCGCCCAAAAGGGGCGCAATCAACCAGAGTGTACATTAATGTGGACAGG	3660	4741	ATCCCACTTTCACCATTTGAGACGACGACCGTGCCTCCCAAGACGCGGTGTTCGCGCTCGCAAC	4800
Qy	3661	ACCTCGTGGCTGGCAGGCGCCCGCGGGCGCGCTCCATGACACCATGACGTGTGGCA	3720	4741	ATCCCACTTTCACCATTTGAGACGACGACCGTGCCTCAAGACGCGGTGTTCGCGCTCGCAGC	4800
Db	3661	ACCTCGTGGCTGGCAGGCGCCCGCGGGCGGTTCTTGACACCATGACCTGTGGCA	3720	4801	GGCGAGGTAGAACTGCGACGGGTAGGAGTGGCACTACAGGTTTGTGACTCCAGGAGAAC	4860
Qy	3721	GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCAATCCGGTGGCGCGCGGAGGCG	3780	4801	GGCGGGTTAGGACTGCGACGGGTAGGAGGAGCATCTACAGTTTGTGACTCCGGGAGAAC	4860
Db	3721	GCTCAGACCTTTACTTGGTCAAGACATGCTGACGTCAATCCGGTGGCGCGCGGCGG	3780	4861	GGCCCTCGGGCATGTTCCGATTCCTCGTCTCTGTGAGTGTCTATGACGCGGGTGTGCTT	4920
Qy	3781	ACAGCAGGGGAGTCTACTCTCCCGCAGGCGCGTCTCTACCTGAAAGGCTCTCGGGTG	3840	4861	GGCCCTCGGGCATGTTCCGATTCCTCGTCTCTGTGAGTGTCTATGACGCGGGTGTGCTT	4920
Db	3781	ACAGTAGGGGAGCCTGCTCTCCCGCAGGCGCTGTCTCTACTTTGAAGGGCTCTTCGGGTG	3840	4921	GGTATGAGCTCAGCGCGCTGAGACCTCGGTAGGTTGCGGGCTTACCTAAATACACAG	4980
Qy	3841	GTCCATTTGCTTTCCTTTCGGGCGACGTCGTGGGCGCTTTCGGGCTGTGTGTGACACC	3900	4921	GGTACGAGCTCACCCCGCGAGACCTCGGTAGGTTGCGGGCTTACCTGAACACACAG	4980
Db	3841	GTCCATGCTCTCGCCCTTCGGGCGACGTCGTGGGCGCTTTCGGGCTGTGTGTGACACC	3900	4981	GGTTGCCGCTGCGCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGCTCACCC	5040
Qy	3901	GGGGGTGCGAAGCGGTGGACTTTCATACCCGTTGAGTCTATGGAACTACCATCGGCT	3960	4981	GGTTGCCGCTTTCGCCAGGACCATCTGGAGTTCTGGGAGAGTGTCTTACAGGCTCACCC	5040
Db	3901	GGGGGTGCGAAGCGGTGGACTTTCATACCCGTTGAGTCTATGGAACTACCATCGGCT	3960	5041	ACATAGATGCCCACTTTCCTGTCCAGACTAAACAGGACGAGAGACAACTTTCCTTACCTGG	5100
Qy	3961	CTCGGCTTTCACAGACACTCAACCCCGCGCTGTACCGCAGACATTCGAAGTGGCAC	4020	5041	ACATAGATGCACTTCTTGTCTCCAGACCAAGCAGGACGAGAGACAACTTCCCTTACCTGG	5100
Db	3961	CTCGGCTTTCACAGACACTCAACCCCGCGCTGTACCGCAGACATTCGAAGTGGCAC	4020	5101	TGGCATATCAAGTACAGTGTGCGCAGGGCTCAAGTCTCACTCCATCTGATGCGGACAAA	5160
Qy	4021	ATCTGACAGCTCTACTTGGCAGCGGCAAGACACAAAGTGCCTGCGCTGTATGACGCC	4080	5101	TAGCATACCAAGCCACGCTGTGCGCAGGGCTCAGGCCCACTCCATCATATGGGATCAAA	5160
Db	4021	ACCTACAGCTCCACTTGGCAGCGGCAAGAGTACTAAAGTGCCTGCGCTGTATGACGCC	4080	5161	TGTGGAAGTGTCTCATACGGCTGAAACCTACATCTGACGCGGCGCAACCCCTGCTGTATA	5220
Qy	4081	AAGGCTACAAAGTGTCTCTGAACCCGTCGCTGCGCGCACTTATAGGTTTGGGCGT	4140	5161	TGTGGAAGTGTCTCATACGGCTGAAACCTACATCTGACGCGGCGCAACCCCTGCTGTATA	5220
Db	4081	AAGGCTACAAAGTGTCTCTGAACCCGTCGCTGCGCGCACTTATAGGTTTGGGCGT	4140	5221	GGCTAGGAGCGCTCCAAAATGAGGTCACTCTCACAACCCCACTAACTAAATACATCATGG	5280
Qy	4141	ATATGTCCAAGGACACGGTATCGACCCCTAACATCAGAACTGGGTAAAGCAATTACCA	4200	5221	GGCTAGGAGCGCTCCAAAATGAGGTCACTCTCACAACCCCACTAACTAAATACATCATGG	5280
Db	4141	ATATGTCCAAGGACACGGTATCGACCCCTAACATCAGAACTGGGTAAAGCAATTACCA	4200	5281	CATGCAATGTCGGTGACCTGGAGTGTCTACCTAGCAGCTGGTGTGGGCGGAGTCC	5340
Qy	4201	CGGCGGCTCCATTTACGTACTCCACCTATGGCAAGTTCTTGGCGAGCGGTGTTCTG	4260	5341	TTGCACTTTGGCGGCACTGCTGACGACGACGACGCTGATTTGTTGGGAGGATCA	5400
Db	4201	CAGCGGCGGCTCAGATATCTACCTATGGAAGTTCTTGGCGAGTGGTGTGCTG	4260	5341	TTGCACTTTGGCGGCACTGCTGACGACGACGACGCTGATTTGTTGGGAGGATTA	5400
Qy	4261	GGGCGCTATGACATCAATATGTGATGAGTGCCACTCAACTGACTCGACTACCATCT	4320	5401	TCCTGTCGGGAAAGCCAGCTGTCTTCCGACAGGGAAGTCTCTACAGGAGTTCGATG	5460
Db	4261	GGGCGCTATGACATCAATATGTGATGAGTGCCACTCAACTGACTCGACTACCATCT	4320	5401	TCCTGTCGGGAGGCGCGCTATTTGTTCCGACAGGAGCTTCTCTACAGGAGTTCGATG	5460
Qy	4321	TGGGCATCGGACATGCTGGAACGAGGAGAGCGCTGAGCGCGCTGCTGCTGCTG	4380	5461	AGATGGAAGAGTGTGCTTCAAACTTCTTACATCGAGCAGGGAATGAGCTGCCGAGC	5520
Db	4321	TGGGCATCGGACATGCTGGAACGAGGAGAGCGCTGAGCGCGCTGCTGCTGCTG	4380	5461	AAATGGAAGAGTGTGCTTCAAACTTCTTACATCGAGCAGGGAATGAGCTGCCGAGC	5520
Qy	4381	CCACCGCTACACCTCCGGGATCGGTACCGTGCCACACCCCAATATCGAGGAATAGGCC	4440	5521	AAATGGAAGAGTGTGCTTCAAACTTCTTACATCGAGCAGGGAATGAGCTGCCGAGC	5520
Db	4381	CCACCGCTACACCTCCGGGATCGGTACCGTGCCACACCCCAATATCGAGGAATAGGCC	4440	5521	AAATGGAAGAGTGTGCTTCAAACTTCTTACATCGAGCAGGGAATGAGCTGCCGAGC	5520
Qy	4441	TGTCACCAATGAGAGATCCCTTCTATGCAAGGCAATCCCATTTGAGGCGCATCAAGG	4500	5521	AAATGGAAGAGTGTGCTTCAAACTTCTTACATCGAGCAGGGAATGAGCTGCCGAGC	5520
Db	4441	TGTCACCAATGAGAGATCCCTTCTATGCAAGGCAATCCCATTTGAGGCGCATCAAGG	4500	5581	CTCCCGTGTGGAGTCCAAAGTGGCGAGCCCTTGGAGACCTTCTGCGGAGGACATGTTGGA	5640
Qy	4501	GGGAGGAGCATCTCATTTTCTGTCATTCGAAGAGAGTGGCAGCTCCCGCAGAACG	4560	5581	CTCCCGTGTGGAGTCCAAAGTGGCGAGCCCTTGGAGACCTTCTGCGGAGGACATGTTGGA	5640
Db	4501	GGGAGGAGCATCTCATTTTCTGTCATTCGAAGAGAGTGGCAGCTCCCGCAGAACG	4560	5641	ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGGAAACCCCGGA	5700
Qy	4561	TGACAGGCTCGGACTGACATGATTTACCGGGCTTGTGATGTGCTGCTCATAC	4620	5641	ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGGAAACCCCGGA	5700
Db	4561	TGACAGGCTCGGACTGACATGATTTACCGGGCTTGTGATGTGCTGCTCATAC	4620	5701	TAGCATCATTTAGTGGCATTTACGCTTCTATCTAGCCCGCTCACACCCCAAAACACCC	5760
Qy	4621	CGCTATCGGAGAGCTGCTTGTGCGCAACAGACGCTCTAATACGCGGTTTACCGGCG	4680	5701	TAGCATCATTTAGTGGCATTTACGCTTCTATCTAGCCCGCTCACACCCCAAAAGTACCC	5760
Db	4621	CAACTATCGGAGAGCTGCTTGTGCGCAACAGACGCTCTAATACGCGGCTTACCGGCG	4680	5761	TCCTGTTTACATCTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGCT	5820
Qy	4681	ATTTTGAATCAGTATCGACTGCAATATCATGTGTCAACCCAGACAGTGCATTCAGTTGG	4740	5761	TCCTGTTTAAACATCTTGGGGGGTGGGTGGCTGCCAACTCGCTCTCCAGCGCTGCT	5820

QY 5821 CAGCTTTTCGTGGCGCCGCGCATCGCGGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG 5880
DB |||||
5821 CGGCTTTTCGTGGCGCGCGCATCGCGGTGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG 5880
QY 5881 TGCTCGTGGAACATCTTGGCGGGCTATGGGGCAGGGGTAGCGGCGCAGCTCGTGGGCTTTA 5940
DB |||||
5881 TGCTTGTGGACATCTGCGGGTATGGAGCAGGAGTGGCGGCGGCTCGTGGGCTTTA 5940
QY 5941 AGGTCAATGAGCGGCGAGGTCCTCCACCGAGGACCTGGTCAACTTACTCCCTGCCATCC 6000
DB |||||
5941 AGGTCAATGAGCGGCGAGATGCCCTCTACCGAGGACCTGGTCAATCTACTTCTCGGCATCC 6000
QY 6001 TCTCTCTCGTGTGCCCTGCTGCTCGGGTCTGTGCGCAGCAATACTTGGCTGGCAGCTGG 6060
DB |||||
6001 TCTCTCTCGGCGCCTCGTGTGCTCGGGTCTGTGTGCAGCAATACTTGGCTGGCAGCTGG 6060
QY 6061 GCCGGGAGAGGGGCTGTGCAGTGATGAACCGGCTGATAGCGTTTCGTCGGGGTA 6120
DB |||||
6061 GTCCGGAGAGGGGCTGTGCAGTGATGAACCGGCTGATAGCGTTTCGCTCGCGGGTA 6120
QY 6121 ACCAGTCTCCCTACGCATATGTCCTGAGCGAGCGCTGCAGCAGTGTCACCTCAGA 6180
DB |||||
6121 ATCAGGTTTCCCCACGCAATATGCTGTGAGCGAGCGCCGCGAGCGGTGTACTCAGA 6180
QY 6181 TCTCTCTAGCCTTACCATCACTCAACTGCTGAAGCGGCTCCACAGTGGATTAATGAGG 6240
DB |||||
6181 TCCTCTCCAGCCTTACCATCACTCAGCTGCTGAAAGGCTCCACAGTGGATTAATGAGG 6240
QY 6241 ACTGCTCTACGCCATGCTCCGCTCGTGGCTAAGGATGTTTGGGATGGATATGACGG 6300
DB |||||
6241 ACTGTCCACACCGTGTCCGGCTCGTGGCTAAGGATGTTTGGGACTGGATATGACGG 6300
QY 6301 TGTTGACTACTTCAAGACTGGCTCCAGTCCAACTCTGCGCGGTTTACCGGAGTCC 6360
DB |||||
6301 TGTTGACTACTTCAAGACTGGCTCCAGTCCAAAGCTCTGCGCGAGCTACCGGAGTCC 6360
QY 6361 CTTTCTCTGTCATGCAACCGGGTCAAGGGAGTCTGGCGGGGACGGCATCATGCAAA 6420
DB |||||
6361 CTTTTTCTCTGTCGCAACCGGGGTACAGGAGTCTGGCGGGGAGNCGGCATATGCANA 6420
QY 6421 CCACCTGCCATGCGGAGCACAGATGCGCGGACATGTCAAAAAACGGTTCCATGAGGATCG 6480
DB |||||
6421 CCACCTGCCATGTTGGAGCACAGATCACCGGACATGTCAAAAAACGGTTCCATGAGGATCG 6480
QY 6481 TAGGGCTAGAACCTTCAGCAACACGTGGCACGGACGTTCCCATCAACGCGCATACCA 6540
DB |||||
6481 TCGGGCCTAAGACCTTCAGCAACACGTGGCATGGAACTTCCCATCAACGCGCATACCA 6540
QY 6541 CGGGACCTTGCACACCTCCCGCGGCCAACTATTCCAGGGCGCTATGGCGGGTGGCTG 6600
DB |||||
6541 CGGGCCCTTGACACCTCTCCAGCGCCAACTATTCTAGGGGCGCTGTGGCGGGTGGCGG 6600
QY 6601 CTGAGGAGTACGTGGAGTTACCGGTGCGGGATTTCCACTACGTGACGGGCGATGACCA 6660
DB |||||
6601 CTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACTACGTGACGGGCGATGACCA 6660
QY 6661 CTGACNAAGTAAAGTCCCATGCCAGGTTCCGGCCCCCGAATTTCTTACGGAGGTGGATG 6720
DB |||||
6661 CTGACNAAGTAAAGTCCCATGCCAGGTTCCGGCCCCCGAATTTCTTCTCGGAGGTGGACG 6720
QY 6721 GAGTGGGTTGCACAGGTACGCTCCGGCGTGCAAACTCTTCTACGGGAGGACGTCACCT 6780
DB |||||
6721 GAGTGGGTTGCACAGGTACGCTCCGGCGTGCAAGGCTCTCTCTACGGGAGGAGTTACAT 6780
QY 6781 TCCAGTTCGGGCTCAACCAATACTTGGTTCGGGTGCGAGCTCCCATCGAGCCCGAACCGG 6840
DB |||||
6781 TCCAGTTCGGGCTCAACCAATACCTGTTGGGTTCACAGCTACCATCGGAGCCCGAACCGG 6840
QY 6841 AGTAAAGTGTCTTACTTCACTGCTCACCGATCCCTCCACATTACAGCAGAGACGGCTA 6900
DB |||||
6841 ATGTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATCAACAGCAGAAACGGCTA 6900

QY 6901 AGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGTAGCCAGTTGT 6960
DB |||||
6901 AGCGTAGGTTGGCCAGGGGTCTCCCCCTCTCTTGGCCAGCTCTTTCAGTAGCCAGTTGT 6960
QY 6961 CTGGCCTTCTTTGAAAGGACATGCACTACCCACCATGACTCCCCGGACGCTGACCTCA 7020
DB |||||
6961 CTGGCCTTCTTTGAAAGGACATGCACTACCCACCATGCTCTCTCGGACGCTGACCTCA 7020
QY 7021 TCGAGGCCAACTCTTGTGGCGGAGAGATGGGCGGAAACATCACTCGCTGGAGTCAG 7080
DB |||||
7021 TCGAGGCCAACTCTCTGTGGCGGAGAGATGGGCGGGAACATCACCCGCTGGAGTCGG 7080
QY 7081 AGAATAAGGTAGTAATCTGCACTCTTTCGAACCCCTTCAACCGAGGGGATGAGAGGG 7140
DB |||||
7081 AGAACAGGTGTAGTCTCGGACTCTTTTCGACCCCTTTCGAGCGGAGGAGATGAGAGGG 7140
QY 7141 AGATATCCGTCCGCGCGAGATCCGCGAAATCCAGGAAGTTCCCTCAGCGTTGCCCA 7200
DB |||||
7141 AAGTATCCGTTCCGCGCGAGATCCGCGAAATCCAGGAAGTTCCCGCAGCGATGCCCA 7200
QY 7201 TATGGGCAACCGCGACTACAACTCTCCACTGCTAGAGTCTCTGGAAGGACCCCGGACTAG 7260
DB |||||
7201 TCTGGCGCGCCGGATTACAACTCTCACTGTTAGAGTCTCTGGAAGGACCCCGGACTAG 7260
QY 7261 TCCCTCCGGTGGTACACGGATGCCCATTTGCCACCTTACCAAGGCTCTCTCAATACCACTC 7320
DB |||||
7261 TCCCTCCGGTGGTACACGGGTGCCGTTTGCCACCTTATCAAGGCCCTCTCAATACCACTC 7320
QY 7321 CACGAGAAAGAGGACGGTGTCTGCAGAGATCCAAATGTGTCTTCTGCTTTGGCGGAGC 7380
DB |||||
7321 CACGAGAAAGAGGACGGTGTCTTAAACAGAGTCTCTCGTGTCTTCTGCTTTAGCGGAGC 7380
QY 7381 TCGCCACTAAGACCTTCCGCTAGTCCGGATCGTGGCGGTTGATAGCGGACACGGCGACCG 7440
DB |||||
7381 TCGCTACTAAGACCTTCCGAGCTCCGAAATCATCGGCGCTGCACAGCGGACACGGCGACCG 7440
QY 7441 CCCTTCTGACCTGGCCTCCGACGACGGTGACAAAGGATCCGACGTTGAGTCTGCTCT 7500
DB |||||
7441 CCCTTCTGACCGAGCCTCCGACGACGGTGACAAAGGATCCGACGTTGAGTCTGCTCT 7500
QY 7501 CCATCCCCCTTGAAGGGGAGCCGGGGACCCCGATCTCAGACGGGCTCTTGGTCTA 7560
DB |||||
7501 CCATCCCCCTTGAAGGGGAGCCGGGGACCCCGATCTCAGTACGCGGTCTTGGTCTA 7560
QY 7561 CCGTCAAGTGAAGGAGCTAGTGAAGATGTCTGTCTGTCTCAATGTCTTATAGTGACAG 7620
DB |||||
7561 CCGTGAAGGAGGAGCTAGTGAAGATGTCTGTCTGTCTCAATGTCTTACATGACAG 7620
QY 7621 GCGCCTGATCAGCCATGCGCTGCGGAGGAAAGTAAGTGGCCCATCAACCGGTTGAGCA 7680
DB |||||
7621 GCGCCTTATCAGCCATGCGCTGCGGAGGAAAGCAAGTGGCCCATCAACCGGTTGAGCA 7680
QY 7681 ACTCTTGTCTGCTCACCACACATGCTACGCGCACACATCCCGCAGCGCAAGCCTCC 7740
DB |||||
7681 ACTCTTGTCTGCGCCACCACATGCTGCTGTTATGCGCACAACTCTCGCAGCGGAGCCTGC 7740
QY 7741 GGCAGAAAGGTACCTTTTGACAGATTGCCAAGTCTCGATGATCATATTACCGGAGCGTAC 7800
DB |||||
7741 GGCAGAAAGGTACCTTTTGACAGATTGCCAAGTCTCGATGATCATATTACCGGAGCGTAC 7800
QY 7801 TCAAGGAGATGAAGCGAGGCGTCCACAGTTAAGGCTTAAGCTTATCTATATAGAGGAGG 7860
DB |||||
7801 TCAAGGAGATGAAGCGAGGCGTCCACAGTTAAGGCTAAGCTTATCCGTATCCGTAGAGGAG 7860
QY 7861 CTTGAAGCTGAGCGGCCACATTTGGCCAAATCCAAATTTGGCTATGGGCGAAAGGAGC 7920
DB |||||
7861 CTTGAAGCTGAGCGGCCACATTTGGCCAAATCCAAATTTGGCTATGGGCGAAAGGAGC 7920
QY 7921 TCCGGAACCTATCCAGCAGGCGGCTTAACCAATCCGCTCCGTGTGGGAGGACTTGTCTGG 7980
DB |||||
7921 TCCGGAACCTATCCAGCAGGCGGCTTAACCAATCCACTCCGTGTGGGAGGAGACTTGTCTGG 7980
QY 7981 AAGACACTGAAACACCAATTTGACACCAATCATGCGCAAAAGATGAGGTTTTTCTGCGTCC 8040

1981 GCAACTGGTTGGCTGTAATGATGAATAGTACTGGTTCACTAAGACGTGCGGAGGTC 2040
1983 GCAACTGGTTGGCTGTAATGATGAATAGTACTGGTTCACTAAGACGTGCGGAGGTC 2042
2041 CCCCGGTGAACATCGGGGGGTCGGTAAACCGACCTTGATCTGCCCCACGAGCTGCTTCC 2100
2043 CCCCGGTGAACATCGGGGGGTCGGTAAACCGACCTTGATCTGCCCCACGAGCTGCTTCC 2102
2101 GGAAGCACCCCGAGGCTACTTACACAAATGTGGCTCGGGCCCTGGTTGACACCTAGGT 2160
2103 GGAAGCACCCCGAGGCTACTTACACAAATGTGGCTCGGGCCCTGGTTGACACCTAGGT 2162
2161 GCTTAGTAGACTACCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCT 2220
2163 GCTTGGTAGACTACCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCT 2222
2221 TTAAGGTTAGGATGATGTAGGGGGGTCGGACACAGGCTCAATGCCGATGCAATTTGGA 2280
2223 TCAAGGTTAGGATGATGTAGGGGGGTCGGACACAGGCTCGATGCTGCAATTTGGA 2282
2281 CTCGAGGAGCGCTGTAACTTGGAGGACAGGATAGGTGAGAACTCAGCCCGCTGCTGC 2340
2283 CCCGAGGGAACTTGCACCTTGGAGGATAGGATAGTTGGAGCTCAGCCCGCTACTGC 2342
2341 TGTCTACACAGAGTGGCAGATACTGCCCTGTGCTTTTCAACCCCTACCGGCTTTATCCA 2400
2343 TGTCTACACAGAGTGGCAGGCTGCTGCCCTGTTCTTTCACACCCCTACCGGCTCTGTCCA 2402
2401 CTGGTTTGTATCTCCATCAGAACATCGTGAGCGTGCATACCTCTAGCGGTGAGGT 2460
2403 CTGGTTTGTATCTCCATCAGAACATCGTGAGCGTGCATACCTCTAGCGGTGAGGT 2462
2461 CAGCGTTTGTCTCTTGTGCAATCAATGGGAGTACATCTGTTGCTTTTCTCTCTG 2520
2463 CCGCAGTTGTTTCTTGTGCAATCAATGGGAGTACATCTGTTGCTTTTCTCTCTG 2522
2521 CAGACGCGCGGTGTGCTGCTTGTGATGATGCTGCTGATAGCCAGGCTGAGCGG 2580
2523 CCGACGCGCGGTGTGCTGCTTGTGATGATGCTGCTGATAGCCAGGCTGAGCGG 2582
2581 CTTAGAGAACTTGGTGGTCTCAATCGGGGTCGGTGGCGGAGCCATGTTCTCT 2640
2583 CTTAGAAAACTGGTGGTCTCAATCGGGGTCGGTGGCGGAGCCATGTTCTCT 2642
2641 CTTTCTTGTGTTCTCTGCGCGCTGTGATCAATTAAGGCGAGGCTGCTCTCGGCGG 2700
2643 CTTTCTTGTGTTCTCTGCGCGCTGTGATCAATTAAGGCGAGCTGTTCCCGGGGAG 2702
2701 CGTATGCTTTTATGGGCTATGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
2703 CATATGCTTTCTATGGAGTATGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2762
2761 CTTACGCTTTGGAACGGGAGATGGCTGATCGTGGGGGGTGGGTTCTGTAGTCTGG 2820
2763 CTTACGCTTTGGAACGGGAGATGGCTGATCGTGGGGGGTGGGTTCTGTAGTCTGG 2822
2821 TATTTCTTGACCTTCTCACCATACTACAAGTGTCTCTACCTAGGCTCATATGTTGTTAC 2880
2823 TACTCTTGACTTTGTACCATACTATATAAGAGTTCCTCGCCAGGCTCATATGTTGTTGC 2882
2881 AATACTTTATCACCAGAGCCGAGCGCACATGCAAGTGTGGGTCCCCCCCCCTCAACGTT 2940
2883 AATACTTTATCACCAGAGCCGAGCGCACCTGCAAGTGTGGATCCCCCCCCCTCAACATTC 2942
2941 GGGAGGCGCGATGCCCATATCTCTCTCAAGTGTGGGTTCTATCGAGTTAATTTTGG 3000
2943 GGGGGGCGCGATGCCCATATCTCTCTCGGCTGTAGTCCACCCAGAGCTAATCTTTG 3002
3001 ACATCAACAACTCTGCTCGCCATCTCTCGGCGGCTCATGTTGCTCGAGCTGGCATAA 3060
3003 ACATCAACAACTCTGCTCGCCATCTCTCGGCTCATGTTGCTCGAGCTAGCATAA 3062

3061 CGAGAGTCCGCTACTTTCGTGCGGCTCAAGGGCTCATTCGTGATGCATGTTAGTGCAG 3120
3063 CTCAGTGGCGCTACTTTCGTACGCGCCCAAGGGCTCATTCGTGATGCATGTTAGTGCAG 3122
3121 AAGTCCCGCGGGCTCAATATGTCCAAATGGTCTTCAATGAAGCTGGGCGGCTCAACAGGTA 3180
3123 AGGTAGCCGGGGCCATATATGTCCAAATGGCTTGTGAAGCTGACCGCATGACAGGTA 3182
3181 CGTACGTTTATTAACCATCTTACCCACTGGGACTGGGCCCAAGGGCTCAACAGGTA 3240
3183 CGTACGTTTATGAACCATCTAATCTCCACTGGGACTGGGCCCAAGGGCTCAACAGGTA 3242
3241 TTGCGGTGGCGGTAGAGCCGCTGCTCTCTCCGCAATGGAGCAAAAGGTCAACCTGGG 3300
3243 TCGCGGTGGAGTAGAGCCGCTGCTCTCTGACATGGAGCAAAAGGTCAACCTGGG 3302
3301 GAGCAGACCCGCTGCTGTGGGACATCATCTTGGGTCTACCCGCTCTCGGCCGAAGGG 3360
3303 GGGCAGACCCGCTGCTGTGGGACATCATCTTGGGTCTACCCGCTCTCGGCCGAAGGG 3362
3361 GGAAGAGATATTTTGGGACCCGCTGATAGTCTCGAAGGGCAAGGGTGGGACTCTCTG 3420
3363 GTAGGAGATACTTCTGGGCGCGCGGATAGTCTTGAAGGGCAAGGGTGGGCTCTCTG 3422
3421 CGCCCATCAGGCTCTCTCCAAACAGCGGGCGTACTTGGTTGGATCATCACTAGCC 3480
3423 CTCCTCATCAGGCTATTTCCAAACAGCGGGCGCTACTTGGTTGGATCATCACTAGCC 3482
3481 TCACAGCCCGGCAAGAACACAGGTCGAAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
3483 TCACAGCCCGGCAAGAACACAGGTCGAAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3542
3541 AATCTTTCTGGGACCTGATCAACCGCGTGTGGACTGTCTACCATGGCGCTGGCT 3600
3543 AATCTTTCTGGGACCTGATCAATGGCGCTGTGGACTGTCTCCATGGTGGCGGCT 3602
3601 CGAAGACCTTAGCGGTCGAAAGGTCGAAAGGTCGAAAGGTCGAAAGGTCGAAAGGTCG 3660
3603 CAAAGACCTTAGCTGGCGCCCAAGGTCGAAAGGTCGAAAGGTCGAAAGGTCGAAAGGTCG 3662
3661 ACCTCGTGGCTGGCGGCG 3720
3663 ACCTCGTGGCTGGCGGCG 3722
3721 GCTCGGACCTTTACTTGGTCAAGAGATGCTGATGCTATTCGCGGTGGCGCGCGCGCGCG 3780
3723 GCTCAGACCTTTATTTGGTCAAGAGATGCTGATGCTATTCGCGGTGGCGCGCGCGCGCG 3782
3781 ACAGAGGGGAGTCTACTCTCTCCAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3840
3783 ACAGTGGGGAAGCTACT 3842
3841 GTCCATTTGCTTTGCGCTTCGCGGCGAGCTGCTGGGCGCTCTCTCTCTCTCTCTCTCTCTCT 3900
3843 GTCCGCTGCTCTGCGCTTCGAGGACGCTGTGGGCACTCTCTCTCTCTCTCTCTCTCTCTCT 3902
3901 GGGGGGTGCGGAAGCGGCTGGAATTTATACCCGTTGAATCAATGGAACTACTATGCGGT 3960
3903 GGGGGGTGCGGAAGCGGCTGGAATTTATACCCGTTGAATCAATGGAACTACTATGCGGT 3962
3961 CTCGCGTCTTTCAGAGCACTCAACCCCGCGCTGTACCGGAGCAATTCGAGTGGCAC 4020
3963 CTCGCGTCTTTCAGAGCACTCAACCCCGCGCTGTACCGGAGCAATTCGAGTGGCAC 4022
4021 ATCTGACGCTCTTACTGGGAGCGGCAAGAGCAACAAAGTCCGCGCTGCTATGCGAGCC 4080
4023 ATCTGACGCTCTTACTGGGAGCGGCAAGAGCACTAAGTGGCGGCTGCTATGCGAGCC 4082
4081 AAGGTTCAAGGTCGTCTCTGAAACCGCTCGGTTGCCGCACTTACGTTTGGGGGT 4140
4083 AAGGTTCAAGGTCGTCTCTGAAACCGCTCGGTTGCCGCTACCTTGGGTTTGGGGGT 4142
4141 ATATGTCCAAGGCACACGGTATCGACCCCTAATCATCAGAACTGGGGTATAGGACCATATCA 4200

||||| 4143 ATATGCTAAAGCACAATGGTATCGACCCCAACATCAGAACTGGGGTAAGGGCCATCACCA 4202
|||||
||||| 4201 CGGGGGCTCCATTACGTACTCCACCTATGGCAAGTTCCTTGGCCGACGGTGGCTGTTCTG 4260
|||||
||||| 4203 CGGGGGCCCTATTACATACTCCACCTACGGCAAGTTCCTTGGCCGACGGGTTGTTCCG 4262
|||||
||||| 4261 GGGGGCTCTATGACATCAATAATGATGATGAGTGCCACTCAACTGACTCGACTACCATCT 4320
|||||
||||| 4263 GGGGGCCCTATGACATCAATAATGATGATGAGTGCCACTCAACTGACTCGACTACCATCT 4322
|||||
||||| 4321 TGGGCATCGGCACAGTCTCTGGAACCAAGCGGAGACGGCTGGAGCGGGCTCGTGGTCTG 4380
|||||
||||| 4323 TGGGCATTGGCACAGTCTCTGGACCAAGCGGAGACGGCTGGAGCGGGCTCGTGGTCTG 4382
|||||
||||| 4381 CCACCGCTACACCTCGGGATCGGTTACCGTGCCACACCCCAATATCAGGAAATAGGCC 4440
|||||
||||| 4383 CCACCGCTACCGCTCGGGATCGGTCACCGTGCCACACCCCAATATGAGGAGGTGGGCC 4442
|||||
||||| 4441 TGTCCAAACAATGGAGAGATCCCTTCTATGGCAAGCCATCCCAATTGAGGCCATCAAGG 4500
|||||
||||| 4443 TGTCCAAACGCTGGAGAAATCCCTTCTACGGCAAGCCATCCCAATTGAGGTCAACAAG 4502
|||||
||||| 4501 GGGGGAGGCATCTCATTTTCTGCCATTCCAAAGAAGAAATGTACGAGCTCGCGCAAAAGC 4560
|||||
||||| 4503 GGGGAAGACATCTCATTTCTGCCATTCCAAGAAGATATGACGAGCTCGCGCAAAAGC 4562
|||||
||||| 4561 TGACAGGCTCGGACTGAACGCTGTAGCATATTAACGGGGCTTGAATGTCCGTCAATAC 4620
|||||
||||| 4563 TATCAGCCCTCGGACTTAATGTGTAGCATATTAATCGGGGTCTTGATGTGTCCGTCAATC 4622
|||||
||||| 4621 CGCCTATCGGAGACGTCGTTGTCGGCAACAGACGCTCTAATGACGGGTTTACCGGGG 4680
|||||
||||| 4623 CGACCAACGGAGACGTCGTTGTCGGTGGCAACAGACGCTCTAATGACGGGCTTTACCGGG 4682
|||||
||||| 4681 ATTTTGACTCAGTGATCGACTGCAATFACATGTGTCAACCCAGACAGTCGACTTCAGCTTGG 4740
|||||
||||| 4683 ACTTTGACTCAGTGATCGACTGTACACATGTGTCAACCCAGACAGTCGATTTACGGCTG 4742
|||||
||||| 4741 ATCCACCTTACCAATTGAGACGACGCGTGCCCAAGACGGGTGTGCGCTCGCAAC 4800
|||||
||||| 4743 ATCCCACTTCACTCAGACGACGACCGTGCCCAAGACGAGTCGGCGCATCACAGC 4802
|||||
||||| 4801 GCGAGGTAGAACTGCGAGGGTACGAGTGGCATCTACAGGTTTGTGACTCCAGAGAAC 4860
|||||
||||| 4803 GCGGGGTAGGACTGTGTAGGGGAGGAGGAGCATCTACAGGTTTGTGACTCCAGAGAAC 4862
|||||
||||| 4861 GGCCTCGGGCATGTTCCGATTCTTCGGTCTGTGTAGTGCTATGACGGGCTGTGCTT 4920
|||||
||||| 4863 GGCCTCGGGCATGTTCCGATTCTTCGGTCTGTGTAGTGCTATGACGGGCTGTGCTT 4922
|||||
||||| 4921 GGTATAGCTACGCGCGCTGAGACCTCGGTTAGGTTGCGGCTTACCTAAATACACAG 4980
|||||
||||| 4923 GGTATAGCTACGCGCTGCTGAGACCTCGGTTAGGTTGCGGCTTACCTGAAATACACCA 4982
|||||
||||| 4981 GGTGCGGCTGCGAGGACCTGTGGAGTCTTGGAGAGCGCTTCAAGGCTCACCC 5040
|||||
||||| 4983 GGTGCGGCTGTGCCAGGACCATCTGGAGTTTGGAGAGCGGTCTCCAGGCTCACCC 5042
|||||
||||| 5041 ACATAGATGCCACTTCTCTGCCAGCTAAACAGGCGAGAGACAATTTCTTACCTGG 5100
|||||
||||| 5043 ACATAGATGCCATTTTCTGTCCCAGACTAACAGGCGAGAGACAATTTCCCTACCTGG 5102
|||||
||||| 5101 TGGCATATCAAGCTACAGTGTGGCCAGGGCTCAAGCTCACTCATCTGTGGGACCAAA 5160
|||||
||||| 5103 TAGCATACCAAGCCACAGTGTGGCCAGAGCTCAAGCTCACCTCATCTGAGATCAAA 5162
|||||
||||| 5161 TGTGGAGTGTCTCATACGGCTGAACCTACACTGACGGGCCAACACCCCTGCTGTATA 5220
|||||
||||| 5163 TGTGGAAGTGTCTCATACGGCTCAAAACCCAGCGCTGACGGGCCAACACCCCTGCTGTATA 5222
|||||
||||| 5221 GGTAGAGCCGCTCCAAAATGAGGTCATCTCTCACACACCCCAATAAATACATCATGG 5280
|||||

Db 5223 GGCTAGGACCGCTCAAAAATGAGATCACCTTCACACACCCCATGACCAAAATTCATCATGG 5282
QY 5281 CATGATGTGGCTGACCTGAGAGTGTGCTACTAGCACCTGGTGTGCTAGGCGGAGTCC 5340
|||||
Db 5283 CATGATGTGGCTGACCTGAGAGTGTGCTACTAGCACCTGGTGTGCTAGGCGGAGTCC 5342
|||||
QY 5341 TTGAGCTTTTGGCCGCTATGCTGCTGACGACGAGGAGTGTGGTCAATTTGGGCGAGATCA 5400
|||||
Db 5343 TTGAGCTCTGGCTGCTATTTGTTGACAAACAGGCAAGTGTGGTCAATTTGGGTAGGATCA 5402
|||||
QY 5401 TCTTGTCCGGAAGGCACTGTGTTCCCGACAGGAAAGTCTCTACAGAGTTTCGATG 5460
|||||
Db 5403 TCTTGTCCGGAAGGCGGCTGTTATTTCCCGACAGGAAAGTCTCTACCGGAGTTTCGATG 5462
|||||
QY 5461 AGATGGAAGAGTGTGCTTCACAACTTCTTATCATCGACAGGGAATGACGCTCGCGGAGC 5520
|||||
Db 5463 AGATGGAAGAGTGTGCTTCACAACTTCTTATCATCGAAACAGGGAATGACGCTTCGCGAGC 5522
|||||
QY 5521 AATTCAAGCAAAAGCGCTCGGGTTTGTGCAAAACGGCCACCAAGCAAGCGGAGGCTGTG 5580
|||||
Db 5523 AATTCAAGCAAGGCGCTCGGATTGCTGCAAAACAGCCACCAAGCAAGCGGAGGCTGTG 5582
|||||
QY 5581 CTCCTGTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACCTTCTGGCGGAAGCACTGTGGA 5640
|||||
Db 5583 CTCCTGTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACCTTCTGGCGGAAGCACTGTGGA 5642
|||||
QY 5641 ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTCGTGGAAACCCCGCA 5700
|||||
Db 5643 ATTTTCATCAGCGGGAATACAGTACCTAGCAGGCTTGTCCACTCTGCTGGAAACCCCGCA 5702
|||||
QY 5701 TAGCATCAATTGATGGCATTTTACAGCTTCTATCACTAGCCCGCTCACCAACCCCAACACCC 5760
|||||
Db 5703 TAGCATCACTGATGGCATTTACAGGCTTCTATCACCAGCCCGCTCTCCACCAAGATACCC 5762
|||||
QY 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAACTGCTCTCTCCAGCGCTGCT 5820
|||||
Db 5763 TATTATTTAAACATCTGCGGGGATGGGTGGCTGCCAACTGCTCTCTCCAGCTGTGCTT 5822
|||||
QY 5821 CAGCTTTCGTGGGCGCGGATCGCGGAGCGGCTTGGGAGCATAGSCCTTGGGAAG 5880
|||||
Db 5823 CGGCTTTCTGCGGCGCGGATCGCGGCTGCGGCTGCGGCTGCGGCTGCTAGCTTTTA 5882
|||||
QY 5881 TGCTCGTGACATCTTGGCGGCTATGGGGGAGGGGTAGCGGCGCATCTGCTGGCTTTTA 5940
|||||
Db 5883 TGCTGTGACATCTTGGCGGATATGGGAGGGGTGCTGCGGCTGCTAGCTTTTA 5942
|||||
QY 5941 AGGTCAATGAGCGGCGAGGTGCCCTCCACGAGGACCTGGTCAACTTCTCTCCCTGCCATCC 6000
|||||
Db 5943 AGATCATGAGCGGCGAGGTGCCCTCCACGAGGACCTGGTTAACTTCTCTCCCTGCCATCC 6002
|||||
QY 6001 TCTCTCTGCTGGCTGCTGCTGGGCTGCTGCGGAGCAATACTGCGTGGCAGCTGG 6060
|||||
Db 6003 TCTCTCCCGGCGGCTGCTGCTGCGGCTGCTGCGGAGCAATACTGCGTGGCAGCTGG 6062
|||||
QY 6061 GCGCGGAGAGGGGCTGTGCAAGTGATGAAACCGGCTGATAGCTTTCGCTTCGCGGGTA 6120
|||||
Db 6063 GCGCGGAGAGGGGCTGTACAGTGATGAAACCGGCTGATAGCTTTCGCTTCGCGGGTA 6122
|||||
QY 6121 ACCAGCTCTCCCTACGCACTATGTCCTGAGAGGACGCTGCGAGCACTGTGTCACTCAGA 6180
|||||
Db 6123 ACCAGCTTCCCGCGCACTATGTCCTGAGAGGACGCTGCGGCGCTGTACTCAGA 6182
|||||
QY 6181 TCCTCTTAGCTTACCATCACTCAACTGCTCAAGGCTCCACAGTGGATTAATGAGG 6240
|||||
Db 6183 TCCTCTCGGCTTACCATCACTCAAGTGGTCTGTAAGGATGTTTGGGACTGATATGACGG 6242
|||||
QY 6241 ACTGCTTACGCGCATGCTCCGGCTCGTGGCTAAGGATGTTTGGGATTTGGATATGACGG 6300
|||||
Db 6243 ACTGCTCCAGCGCATGCTCCGGTCTGTCGCTAAGGATGTTTGGGACTGATATGACGG 6302
|||||
QY 6301 TGTGACTGACTTCAAGACCTGCTCCAGTCAAACTCTCTGCGCGGTTTACCGGGAGTCC 6360
|||||
Db 6303 TGTGACTGACTTCAAGACCTGCTCCAGTCCAGCTCTGCGCGGTTTACCGGGGTTCC 6362
|||||

QY	6361	CTTTCTCTGTCAATCCAAACGCGGTAACAAGGAGTCTGGCGGGGACGGGATCATCAAA	6420
Db	6363	CTTTCTCTCTGTCAACGCGGTAACAAGGAGTCTGGCGGGGACGGGATCATCAGA	6422
QY	6421	CGACCTCCCATCGGAGCAGACATCGCGGACATGTCAAAACGGTCCATCAGGATCG	6480
Db	6423	CCACCTGCCGCTGTGAGCACAATCACCGGACATGTCAAAAACGGTTCATCAGGATCG	6482
QY	6481	TAGGGCTAGAACTCTGCAGCAACACGCTGGCAGGAAAGTTCCCCATCAACGCAATACCA	6540
Db	6483	TCGGGCTAAACCTGCAGCAGCAGCTGGCATGGAAGTTCCCCATCAACGCAATACCA	6542
QY	6541	CGGGAACCTTGCACAACCTCTCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGGTGGCTG	6600
Db	6543	CAGGCCCATCGCAACCTCTCCCGCGCCCAAACTATTCCAGGGCGCTATGGCGGGTGGCGG	6602
QY	6601	CTGAGGAGTACGTGGAGTTACCGGTGTGGGGATTTCCACTACGTGACGGGATGACCA	6660
Db	6603	CTGAGGAGTACGTGGAGTTACCGGGGTGGGGATTTCCACTACGTGACGGGATGACCA	6662
QY	6661	CTGACAACTGAAGTGCATGCCATGCCAGTTCCGGCCCCGAAATTTCTCAGCGAGTGGATG	6720
Db	6663	CTGACAACTGAAGTGCATGCCATGCCAGTTCCGGCCCCGAAATTTCTCAGCTGAGTGGATG	6722
QY	6721	GAGTGGCGTTGCACAGTACGCTCCGGCGTGCAAACCTCTTCTACGGGAGGACGTCACTG	6780
Db	6723	GAGTGGCGTTGCACAGTACGCTCCGGCGTGCAAACCTCTTCTACGGGAGGAGTCACTG	6782
QY	6781	TCAGGTGGGCTCAACCAATATCTTGTGGGTGGAGTCCCATGCGAGCCGGAACCGG	6840
Db	6783	TCCAGGTGGGCTCAACCAATATCTTGTGGGTGGAGTCCCATGCGAGCCGGAACCGG	6842
QY	6841	ACGTAAAGTCTTACTTTCATGCTCACCGATCCCTCCCACTTACAGCAGAGACGCTA	6900
Db	6843	ATGTAGCAGTCTTACTTTCATGCTTACGACCTTACGACCTTCCCATCAGCAGAGAGGCAA	6902
QY	6901	AGCTTAGGCTGGCTAGAGGCTCTCCCGCTCTTTAGCCAGCTCATCAGTACGCGAGTGT	6960
Db	6903	AGCTTAGGCTGGCTAGAGGCTCTCCCGCTCTTTAGCCAGTCTCATCAGTACGCGAGTGT	6962
QY	6961	CTGGCGCTCTTTGAAAGCGACATGCACTACCAACATGACTCCCGGACGCTGACCTCA	7020
Db	6963	CTGGCGCTCTTTGAAAGCGACATGCACTACCAACATGACTCCCGGACGCTTACCTCA	7022
QY	7021	TCGAGGCCAACCCTCTTGTGGCGGAGAGATGGGGGAAACATCACTCGGTGAGTCA	7080
Db	7023	TCGAGGCCAACCCTCTTGTGGCGGAGAGATGGGGGAAACATCACTCGGTGAGTCA	7082
QY	7081	AGAAATAGGTAGTAAATTTGGACTCTTTTCGAAACCGCTTTCAGCGAGGGGATGAGAGG	7140
Db	7083	AGAAATAGGTAGTAAATTTGGACTCTTTTCGAAACCGCTTTCAGCGAGGGGATGAGAGG	7142
QY	7141	AGATATCCGTCGCGCGGAGATCTTCGGAATAATCCAGAAAGTTCCCTCAGCTTGCCCA	7200
Db	7143	AACCATCCGTTGCGCGGAGATCTTCGGAATAATCCAGAAAGTTCCCTCAGCTTGCCCA	7202
QY	7201	TATGGGACGCGCGGACATCAATCTCTCACTGCTAGACTCTGGAAGACCCGACCTACG	7260
Db	7203	TATGGGACGCGCGGATTAACAACCTCTCGTGTGCTAGAGTCTCTGGAAGACCCGACCTACG	7262
QY	7261	TCCCTCCGGTGGTACACGGATGCCATTTGCCACCTTACCAGGCTCTCCAATACCACTC	7320
Db	7263	TCCCTCCGGTGGTACACGGATGCCCTTACCACCTTACCAGGCTCTCCAATACCACTC	7322
QY	7321	CACGAGAAAGAGGACGGTGTCTCTGACAGAAATCCAATGTGTCTCTGCTTGGCGGAGC	7380
Db	7323	CACGAGAAAGAGGACGGTGTCTCTGACAGAGTCCACTGTCTCTGCTTGGCGGAGC	7382
QY	7381	TCGCCACTAAGACCTTTCGGTGTAGTCCGATCGTGGCGGCTTGTATAGCGGACGCGGACCG	7440
Db	7383	TTGCTACTAAGACCTTTCGGTGTAGTCCGCTCGTGGCGGCTGTGACAGCGGACGCGGAACTG	7442

QY	7441	CCCTTCTTACCTGGCTCCGACGCGGTGACAAAGGATCCGACGTTGAGTCTGACTCCT	7500
Db	7443	CTCCTCCCGACGAGGCTTCCGACGCGGACCAAGGATCTGACGTTGAGTCTGATTCCT	7502
QY	7501	CCATGCCCCCTCTTGAAGGGGAGCGCGGGGACCCCGATCTCAGCGACGGGTCTTGTGCTA	7560
Db	7503	CCATGCCCCCTCTTGAAGGGGAGCGCGGGGACCCCGATCTCAGCGACGGGTCTTGTGCTA	7562
QY	7561	CGTGTAGTGAAGGCTAGTGAAGATGTCGTCTGCTGCTCAATGTCTTATACCTGAGACAG	7620
Db	7563	CGTGTAGTGAAGGCTAGTGAAGATGTCGTCTGCTGCTCAATGTCTTATACATGAGACAG	7622
QY	7621	GGCCCCGTATCAGCCATCGCTGCGGAGGAAAGTAAGTGCCTCATCAACCCGTTGAGCA	7680
Db	7623	GGCCCCGTATCAGCCATCGCTGCGGAGGAAAGTAAGTGCCTCATCAACCCGTTGAGCA	7682
QY	7681	ACTCTTTGTCTGCTACCAACATCGTCTACGCCAACATCCCGCAGCGCAAGCTCC	7740
Db	7683	ACTCTTTGTCTGCTACCAACATCGTCTATGCTACACATCCCGCAGCGCAGGCTAC	7742
QY	7741	GGCAGAAAGGTCACTTTTGAAGATTTGCAAGTCTCTGATGATCATTACCGGACGCTAC	7800
Db	7743	GGCAGAAAGGTCACTTTTGAAGATTTGCAAGTCTCTGATGATCATTACCGGACGCTAC	7802
QY	7801	TCAGAGATGAAGCGAGGCGTCCACAGTTAAGCTTAAGCTTAAGCTTCTATCTATAGAGAGG	7860
Db	7803	TCAGAGATGAAGCGAGGCGTCCACAGTTAAGCTTAAGCTTAAGCTTCTATCTATAGAGAGG	7862
QY	7861	CCTGCAAGCTGACGCCCCACATTCGCGCCAAATTCCTGCTATGGCTATGGGCGCAAGACG	7920
Db	7863	CCTGTAAGCTGACGCCCCACATTCGCGCAGATCCAAATTTGGCTATGGGCGCAAGACG	7922
QY	7921	TCGGAACCTTATCAGCAAGGCGGTTAAACACATCCGCTCCGTGTTGGAAGGACTTGTGG	7980
Db	7923	TCGGAACCTTATCAGCAAGGCGGTTAAACACATCCGCTCCGTGTTGGAAGGACTTGTGG	7982
QY	7981	AAGACACTGAAACACCAATTTGACACCACTCATGSCAAAGAGTGAAGTCTTCTGCGTCC	8040
Db	7983	AAGACACTGAGACACCAATTTGACACCACTCATGSCAAAGAGTGAAGTCTTCTGCGTCC	8042
QY	8041	AACAGAGAGGAGGCGGCAAGCCAGCTCGCTTATCGTATTCCTCCAGACCTGGGAGTTC	8100
Db	8043	AACAGAGAGGAGGCGGCAAGCCAGCTCGCTTATCGTATTCCTCCAGACCTGGGAGTTC	8102
QY	8101	GTGTATGCGAAGAGTGGCCCTTTACGAGTGTCTCCACCTTCTCAGGCCGTGATGG	8160
Db	8103	GTGTATGCGAAGAGTGGCCCTTTATGAGTGTCTCCACCTTCTCAGGCCGTGATGG	8162
QY	8161	GCTCCTCATACGGATTTCAATATCTCCCAAGCAGCGGCTCGAGTTCCTGGTGAATACCT	8220
Db	8163	GCTCCTCATACGGATTTCAATATCTCCCAAGCAGCGGCTCGAGTTCCTGGTGAATACCT	8222
QY	8221	GGAAATCAAGAAATGCCCTTATGGCTTCTCATATGACACCGCTGTTTGTGACTCAACGG	8280
Db	8223	GGAAATCAAGAAATGCCCTTATGGCTTCTCATATGACACCGCTGTTTGTGACTCAACGG	8282
QY	8281	TCAGTGAAGTGAATTCGTGTTGAGGAGTCAATTTTACCAATGTTGTGACTTGGCCCCCG	8340
Db	8283	TCAGTGAAGTGAATTCGTGTTGAGGAGTCAATTTTACCAATGTTGTGACTTGGCCCCCG	8342
QY	8341	AGGCAGACAGGCGCATTAAGTTCGTCTACAGAGCGGCTTTTACATCGGGGCTCCCTGACTA	8400
Db	8343	AGGCAGACAGGCGCATTAAGTTCGTCTACAGAGCGGCTTTTACATCGGGGCTCCCTGACTA	8402
QY	8401	ACTCAAAAGGCGCAACTGGGTTATCGCGGTCGCGCAAGTGGGCTGTGAGGACTA	8460
Db	8403	ATTCAAAAGGCGCAACTGGGTTATCGCGGTCGCGCAAGTGGGCTGTGAGGACTA	8462
QY	8461	GCTCGGTAAATACCTTCACTGTTTAAAGGCACTGACAGCCTGTGAGCTGCAAAAGC	8520
Db	8463	GCTCGGTAAATACCTTCACTGTTTAAAGGCACTGACAGCCTGTGAGCTGCAAAAGC	8522
QY	8521	TCCAGGACTGACGATGCTCGTGAAACGAGACGACCTTGTGTTATCTGTGAAAGCGCGG	8580

Db 8523 TCCGGGACTGACGATGCTCGTGAACGAGAGCGACTCGTGTCTCATCTGTGAGAGTGGG 8582
QY 8581 GAAACGAGGATGCGGGGCGCTTACGAGCCTTACGAGGCTATGACTAGGTTATTCG 8640
Db 8583 GAAACGAGGATGAGGCGAACTTACGAGTCTTACGAGGCTATGACTAGGTTATTCG 8642
QY 8641 CCCCCCGGGGATCCGCCCAACAGAAATACGACTGAGCTGATTAACATCATCTTCT 8700
Db 8643 CCCCCCGGGGACCGCCCGACAGAAATACGACTTGGAGCTAATAACATCATCTTCT 8702
QY 8701 CCAATGTGTAGTCGCGACGATGCTTGGCAAAAGGTTATCTACCTCCACCGTGACC 8760
Db 8703 CCAATGTGTGGTGGCGACGATGCTTGGCAAAAGGTTATCTACCTCCACCGTGACC 8762
QY 8761 CCACACCCCTTGCACGGGTGCGTGGGAGACAGCTAGACACACTCAATCAACTCTT 8820
Db 8763 CTTCCACCCCTTGCACGGGTGCGTGGGAGACAGCTAGACACACTCAATCAACTCTT 8822
QY 8821 GGTAGGCAATATCATGTATGCGCCACCTATGGCAGGATGATCTGTGATGACTC 8880
Db 8823 GGTAGGCAATATCATGTATGCGCCACCTATGGCAGGATGATCTGTGATGACTC 8882
QY 8881 ACTTTTCTCCATCTTCTAGTCAAGACAACTTGAAGGCTTGGATTTGTCAGATCT 8940
Db 8883 ATTTCTCTCCATCTTCTAGCCAGGACAACTTGAAGGCTTGGATTTGTCAGATCT 8942
QY 8941 ACGGGCTGTCTACTCCATAGACCACTTGACCTACCTCAGATCAATTGAACACTCCATG 9000
Db 8943 ACGGGCTGTCTACTCCATAGACCACTTGACCTACCTCAGATCAATTGAACACTCCATG 9002
QY 9001 GTCTTAGGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCAT 9060
Db 9003 GTCTTAGGCAATTTACACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCAT 9062
QY 9061 GGCTCAGGAACTTGGGTTACCACTCTTGGCAACCTTGGAGACATCGGGCCAGAGTGCC 9120
Db 9063 GGCTCAGGAACTTGGGTTACCACTCTTGGCAACCTTGGAGACATCGGGCCAGAGTGCC 9122
QY 9121 GGGCTAAGCTACTGTCACGGGGGAGGGCCGACCTTGTGGAGATACCTCTTTAACT 9180
Db 9123 GGGCTAAGCTACTGTCACGGGGGAGGGCTGCGCACTTGTGGTAACTCTTTCACT 9182
QY 9181 GGGCAGTAGGACCAAGCTTAACTCACTCACTCCAGTCCCGCGCTCCAGCTGGACTGT 9240
Db 9183 GGGCAGTAGGACCAAGCTTAACTCACTCACTCCAGTCCCGCGCTCCAGTGGACTGT 9242
QY 9241 CTGGCTGTGCTCGTGTGCTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTGCC 9300
Db 9243 CCAGCTGTGCTCGTGTGCTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTGCC 9302
QY 9301 GACCCGCTGTGCTGTGCTTACTCTACTCTTCTAGGGTAGGCATTTACCTGC 9360
Db 9303 GACCCGCTGTGCTGTGCTTACTCTACTCTTCTAGGGTAGGCATTTACCTGC 9362
QY 9361 TCCCAACCGATCAAGGGGAGTAACTCACTCCAGCCTTAAGCCATTTCTGTTTTTT 9420
Db 9363 TCCCAACCGATCAAGGGGAGTAACTCACTCCAGCCTTAAGCCATTTCTGTTTTTT 9422
QY 9421 TT 9474
Db 9423 TT 9482
QY 9475 -----CCTTTCTTTTCCCTCTTTTAAATGGTGGCTCCATCTTAGCCCTAGTCAAGGC 9526
Db 9483 TTTCTTTGTTT 9542
QY 9527 TAGCTGTGAAGGTCCGTGAGCGGATGACTGCAGAGTGTGATCTAGTGGCTCTCTGC 9586
Db 9543 TAGCTGTGAAGGTCCGTGAGCGGATGACTGCAGAGTGTGATCTAGTGGCTCTCTGC 9602
QY 9587 AGATCATGT 9595
|||||

Db 9603 AGATCATGT 9611
RESULT 9
AAD25517
ID AAD25517 standard; DNA; 9413 BP.
XX
AC AAD25517;
DT 26-MAR-2002 (first entry)
XX
DE Hepatitis C virus complete DNA genome.
XX
KW HCV; hepatitis C virus; cytostatic; cancer; immunosuppressive; virucide;
KW antibacterial; fungicide; protozoicide; antirheumatic; antiinflammatory;
KW antiarthritic; rheumatoid arthritis; neuroprotective; multiple sclerosis;
KW immune response; vasotropic; vaccine; gene therapy; autoimmune disease;
KW vasculitis; ds.
XX
OS Hepatitis C virus.
XX
FN WO200176643-A1.
XX
PD 18-OCT-2001.
PF 06-APR-2001; 2001WO-US011372.
PR 07-APR-2000; 2000US-0195680P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Orson FM, Kinsey BM, Bhogal BS;
XX
DR WPI; 2002-066308/09.
XX
PT Composition for oral delivery of vaccines, comprises expression vector
PT containing antigenic genomic sequence, bound to aggregated protein-
PT polycationic polymer conjugate or suspension.
XX
PS Disclosure; Page 84-87; 145pp; English.
XX
CC The invention relates to a composition comprising an expression vector
CC bound to an aggregated protein-polycationic polymer conjugate or
CC suspension. The expression vector contains a promoter polynucleotide
CC sequence operatively linked to a polynucleotide sequence encoding an
CC antigen which is a fragment of a gene or genome associated with an
CC infectious disease, cancer and autoimmune disease such as rheumatoid
CC arthritis, vasculitis, and multiple sclerosis, pathogenic genomes
CC consisting of bacterium, fungus, protozoa and virus such as human
CC immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C
CC virus (HCV), influenza and respiratory syncytial virus (RSV), and
CC optionally comprising a nucleotide sequence encoding a cytokine (or a
CC cytokine expression vector), is useful for inducing an immune response
CC (systemic and/or mucosal) in an organism. The cytokine expression vector
CC contains a sequence for granulocyte macrophage-colony stimulating factor
CC (GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding
CC the antigen and the cytokine are under transcriptional control of same or
CC different promoter polynucleotide sequences. The expression vector, as a
CC DNA vaccine is useful for treating a condition in an organism. The
CC present sequence is hepatitis C virus complete DNA genome related to the
CC invention
XX
SQ Sequence 9413 BP; 1886 A; 2816 C; 2694 G; 2017 T; 0 U; 0 Other;
Query Match 86.6%; Score 8311.8; DB 6; Length 9413;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 8724; Conservative 0; Mismatches 687; Indels 0; Gaps 0;

QY 14 TGGGGGCGACCTCCACCATGAATCACTCCCTGTGAGAACTACTGTCTTCCACGAGAA 73
|||||
Db 2 TGGGGGCGACCTCCACCATGAATCACTCCCTGTGAGAACTACTGTCTTCCACGAGAA 61
|||||
QY 74 AGCGTCTAGCCATGGCGTTAGTATGATGTCGTGCAGCCTCCAGGACCCCTCCCGG 133
|||||

Db	62	AGCGTCTAGCCATGGCGTTAGTATGATGTTGTGCAGCCTCCAGGACCCCGCTCCCGG	121
Qy	134	AGAGCCATAGTGGTCTGCCGAAACCGGTGAGTACACCGGAATGCGCAGGACGACGGGTCC	193
Db	122	AGAGCCATAGTGGTCTGCCGAAACCGGTGAGTACACCGGAATGCGCAGGACGACGGGTCC	181
Qy	194	TTTCTTGGATCAACCCGGCTCAATGCTGGAGATTTGGGGGTGCCCGCGAGACTGCTAG	253
Db	182	TTTCTTGGATCAACCCGGCTCAATGCTGGAGATTTGGGGGTGCCCGCGAGACTGCTAG	241
Qy	254	CGAGTAGTGTGGGTGCGAAAGGCTTGTGTGCTGCTGCTGATAGGTGCTTGGCAGTG	313
Db	242	CGAGTAGTGTGGGTGCGAAAGGCTTGTGTGCTGCTGCTGATAGGTGCTTGGCAGTG	301
Qy	314	CCCCGGAGGTCTGCTAGACCGGTGACCATGACGACGAATCTTAAACCTCAAAAGAAAC	373
Db	302	CCCCGGAGGTCTGCTAGACCGGTGACCATGACGACGAATCTTAAACCTCAAAAGAAAC	361
Qy	374	CAAACTGTAACCAACCGCCGCCACAGGACGTCAAGTTTCCGGGGGTGGTCAGATCGT	433
Db	362	CAAACTGTAACCAACCGCCGCCACAGGACGTCAAGTTTCCGGGGGTGGTCAGATCGT	421
Qy	434	TGCTGAGATTTACTGTTTCCGCGCAGGGGCCCGAGGTTGGGTGTCGCGCGACTAGGAA	493
Db	422	TGCTGAGATTTACTGTTTCCGCGCAGGGGCCCGAGGTTGGGTGTCGCGCGACTAGGAA	481
Qy	494	GGCTTCCGAGCGGTGCGAACTGCTGGAAGCGGACCACTATCCCAAAGGCTCGCGAGC	553
Db	482	GACTTCCGAGCGGTGCGAACTGCTGGAAGCGGACCACTATCCCAAAGGCTCGCGAGC	541
Qy	554	CGAGGGCAGGCGCTGGGCTCAGCCCGGGTACCTTGGCCCTCTATGGCAATGAGGCGCT	613
Db	542	CGAGGGTAGAACCTGGGCTCAGCCCGGTACCTTGGCCCTCTATGGCAATGAGGCGCT	601
Qy	614	GGGTGGGCAAGATGGCTCTGTCAACCCCGCGCTCCCGGCTAGTTGGGGCCCGACGGA	673
Db	602	GGGTGGGCAAGATGGCTCTGTCAACCCCGGTCTCGGCTAGTTGGGGCCCGACGGA	661
Qy	674	CCCCCGGCTAGTGGCGTAACTTGGGTAGGTCATCGATACCTTACATGCGGCTCGC	733
Db	662	CCCCCGGCTAGTGGCGTAACTTGGGTAGGTCATCGATACCTTACATGCGGCTCGC	721
Qy	734	CGATCTCATGGGTACATTCGGCTGCTCGGCGCCCTAGGGGGCGCTGCCAGGCGCTT	793
Db	722	CGACTCATGGGTACATTCGGCTGCTCGGCGCCCTAGGGGGCGCTGCCAGGCGCTT	781
Qy	794	GGCACACGGGTGTCGGGGTCTGGAGGACGGCGTGAACCTATGCAACAGGGAACTTGC	853
Db	782	GGCACATGGTGTCCGGGTCTGGAGGACGGCGTGAACCTATGCAACAGGGAACTTGC	841
Qy	854	TTGCTCTTTCTATCTTCTCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	913
Db	842	TTGCTCTTTCTATCTTCTCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	901
Qy	914	TTATGAGTGGCAACGCTGCGGATATACATGTACGAACTGCTCCAACTCAAG	973
Db	902	TTAGAGGTGGCAACGCTGCGGATATACATGTACGAACTGCTCCAACTCAAG	961
Qy	974	CATTGTGATGAGGACGCGGCTGATCATGCACTATCCCGGGTGGTGGCTGCTGCTGCT	1033
Db	962	TATTGTGATGAGGACGCGGATCATATGCACTATCCCGGGTGGTGGCTGCTGCTGCT	1021
Qy	1034	GGAGGTAAAGCTCCCTGCTGGGTAGCGCTCATCTCCACGCTCGCGGCGGAGGATGC	1093
Db	1022	GGAGGTAAATTTCTCCCGTGGGTAGCGCTCATCTCCACGCTCGCGGCGGAGGAA	1081
Qy	1094	CAGGTCCCGCTAGCAATAACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1153
Db	1082	CAGATCCCGCTAGCAATAACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1141
Qy	1154	CTGCTCCGCTAGTACGTGGGGATCTCTGGGATCTATTTTCTCGCTCTCCAGCTGTT	1213

Db	1142	CTGTTCCGCTATGATGATGGGATCTCTGCGGATCCGTTTTTCTCGTCTCCAGCTGTT	1201
Qy	1214	CACCTTCTCGCTCGCGCATGAGACACTGAGGACTCAACTGCTCAATCTATCCCGG	1273
Db	1202	CACCTTCTCACCTCGCGGTATGAGACGCTCAAGATTGCAATTGCTCAATCTATCCCGG	1261
Qy	1274	CCATGATCAGGTCAACCGCATGCTGGGATATGATGAACTGCTCACTACACAGC	1333
Db	1262	CCACGTATCAGGTCAACCGCATGCTGGGATATGATGAACTGCTCACTACACAGC	1321
Qy	1334	CCTAGTGGTGTGCGAGTTGCTCCGGATCCCAAGCTGCTGCGGACATGCTGGCGGGG	1393
Db	1322	CCTAGTGGTATCGCAGTACTCCGGATCCCAAGCGCTGCTGCGACATGCTGGCGGGG	1381
Qy	1394	CCACTGGGAGTCTCGCGGGCTTCCCTACTATTCATGCTGAGGAACTGGGCTAAGT	1453
Db	1382	CCACTGGGCTCTAGCGGGCTTCCCTACTATTCATGCTGAGGAACTGGGCTAAGT	1441
Qy	1454	TCGTGATGTCGCTACTCTTTCCCGGCTTACGCTCCCTTTCTCATCTGGGGCGTCTCAGAAAT	1513
Db	1442	CTTGTATGATGCTACTCTTTGCTGGGCTTACGCTCCCTTTCTCATCTGGGGCGTCTCAGAAAT	1501
Qy	1514	GGCGGCGCACACCACTCCGGGTTCAGCTCCCTTTCTCATCTGGGGCGTCTCAGAAAT	1573
Db	1502	GGTAGCTCCAGCACCCAGAGCCTCGTCTCTGCTCTCAAGGCGCCATCTCAGAAAT	1561
Qy	1574	CCAGCTTGTGAATPACCAACGGCAGCTGGCACATCAACAGAGCTGCCCTAAATTTGCAATGA	1633
Db	1562	CCAACTCGTGAACACCAACGGCAGCTGGCACATCAACAGAGCCTCTGAAATTTGCAATGA	1621
Qy	1634	CTCCTTCCAACTGGGTCTTTTCCCGGCTGTTTACGACACAAAGTTCAACTGCTCGG	1693
Db	1622	CTCCTTCCAACTGGGTCTTTTCCCGGCTGTTTACGACACAAAGTTCAACTGCTCGG	1681
Qy	1694	GTCCCGGAGCGCATGGCCAGCTGCGCCCATTTGACTGTTCCGCCAGGGGTGGGGCC	1753
Db	1682	GTCCCGAGCGCATGGCTAGCTGCGCCCATCTGATGAGTTGCTCAGGGGTGGGGTCC	1741
Qy	1754	CATCACCTATPACTAAGCTTAACAGCTCGGATCGAGGCTTATGCTGGGCAATACGGCC	1813
Db	1742	CATCACCTATGATATGCTGAGAGCTCGGACACAGAGGCCATATGCTGGCACTACGGCC	1801
Qy	1814	TCGACCGTGTGGTGTGCTACCCGGTGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1873
Db	1802	TCGACCGTGTGGTGTGCTACCCGGTGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1861
Qy	1874	AAGCCCTGTTGCTGGGGACCAACGATCGTTCGGGTGCTCCCTACGTATAGCTGGGGGA	1933
Db	1862	GAGCCCTGTTGCTGGGGACCAACGATCGTTCGGGTGCTCCCTACGTATAGCTGGGGGA	1921
Qy	1934	GAATGAGACAGACGTGATGCTCCTCAACACACGCGTCCGCAACAGGCACTGCTGCGG	1993
Db	1922	GAATGAGACAGACGTGCTGCTACTTAGCAACACGCGCGGCTCAAGGCACTGCTGCGG	1981
Qy	1994	CTGTACATGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2053
Db	1982	GTGCACCTGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2041
Qy	2054	CGGGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2113
Db	2042	CGGGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2101
Qy	2114	GGCTACTTACACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2173
Db	2102	GGCCACTTACACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2161
Qy	2174	CCCCTACAGGCTTTGGCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2233
Db	2162	CCCCTACAGGCTTTGGCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2221
Qy	2234	GTATGTGGGGGGCTGGAGCACAGCTCAATGCGGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCT	2293
Db	2222	GTATGTGGGGGGCTGGAGCACAGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2281

QY 2294 CTGTAACTTTGGAGGACAGGATAGGTACAGAACTCAGCCGGCTGCTGCTGTCTACAAACA 2353
DB |||||
2282 CTGTGACTTTGGAGGACAGGATAGGTACAGAACTCAGCCGGCTGCTGCTGTCTACAAACA 2341
QY |||||
2354 GTGGCAGATACTGCCCTGTGCTTTTCCACCACCTACCGGCTTTATCCACTGGTTTGATCCA 2413
DB |||||
2342 GTGGCAGATACTGCCCTGTGCTTTTCCACCACCTACCGGCCCTGTCCACTGGCTTGATCCA 2401
QY |||||
2414 TCTCAATCAGAAACATCGTGGAGCGTGCATACCTGTACGGGTGATGGGTACGGTTTGTCTC 2473
DB |||||
2402 TCTTACCGGAAACATCGTGGAGCGTGCATACCTGTACGGTATAGGTCGCGAGTTGTCTC 2461
QY |||||
2474 CTTTGCATCAAAATGGGAGTACATCCTGTGCTTTTCTCTCTCTCTGGCAGACGGCGCGT 2533
DB |||||
2462 CTTTGCATCAAAATGGGAGTATATCCTGTGTCTTTTCTCTCTCTCTGGCGGACGGCGCGT 2521
QY |||||
2534 GTGTGCTCTTGTGGATGATGCTGCTGATAGCCACAGGCTAGAGCGGCTTTAGAGAACTT 2593
DB |||||
2522 CTGTGCTGCTTGTGGATGATGCTGCTGATAGCCACAGGCTGAGGCCACCTTAGAGAACT 2581
QY |||||
2594 GGTGTCTCAATCGGGCTCCGTGGCCGGAGCGCATGATATCTCTCTCTTCTTCTTGTGT 2653
DB |||||
2582 GGTGTCTCAATCGGGCTCTGTGGCCGGAGCGCATGCGCTTCTCTCTCTCTCTCTCTCT 2641
QY |||||
2654 CTTTCTGCGCGCTGCTGCTACATTAAGGGCAGGCTGGCTCTCTGGGCGGCGTATGCTTTTA 2713
DB |||||
2642 CTTTCTGCGCGCTGCTGCTACATAAAGGCGAGGCTGGTCTCTGGGCGGCGATATGCTCTTA 2701
QY |||||
2714 TGGCGTATGGCGCTGCTCTGCTCTACTGCGCTTACCACACGAGCTTAGCCCTTGGA 2773
DB |||||
2702 TGGCGTATGGCGCTGCTCTGCTCTTGTGCGCTTACCACACGAGCTTAGCCATGGA 2761
QY |||||
2774 CCGGAGATGGCTGATCGTGGGGGTGCGGTTCTTGTAGTCTGTGATTTCTTACCTT 2833
DB |||||
2762 CCGAGAGATGGCTGATCGTGGAGGCGCGTCTTGTAGTCTGTGACTCTTGACCTT 2821
QY |||||
2834 GTCCACATACATAAAGTTTCTCACTAGGCTCATATGGTGGTTACAATCTTTATCAC 2893
DB |||||
2822 GTCCACATACATAAAGTTTCTCTGCTAGGCTCATATGGTGGTTACAATATTTTATCAC 2881
QY |||||
2894 CAGAGCCGAGGCGCACATCAAGTGTGGTCCCGCTCAACGTTTCGGGGAGGCGCGA 2953
DB |||||
2882 CAGAGCCGAGGCGCACTTCCAAAGTGTGGTCCCGCTCTCAATGTTTCGGGGAGGCGCGA 2941
QY |||||
2954 TGCATCATCTCTCACTGCTGCGGTTCCATCCAGAGTTAAATTTTGAATCACCAAACT 3013
DB |||||
2942 TGCATCATCTCTTACATGCGCGGTCCATCCAGAGCTAAATCTTTTGACATCACCAAACT 3001
QY |||||
3014 CTTGCTCGGCATACCTGGCCCGCTCATGCTGCTCCAGGCTGGCATAAACGAGTGCCTTA 3073
DB |||||
3002 CTTGCTCGGCATACCTGGCTCCGCTCATGCTGCTCCAGGCTGGCATAACTAGAGTGCCTTA 3061
QY |||||
3074 CTTTGTGCGGCTCAAGGCTCAATCTGCATGCTATGTTAGTGCAGAAAGTTCGCGGGG 3133
DB |||||
3062 CTTTGTACGCGCTCAGGGCTCATCCGTGCATGCTATGTTAGTGCAGAAAGTTCGCTGGAG 3121
QY |||||
3134 TCATTAATGTCCTCAATGGTCTCATGAAGCTGGCGCGCTGACAGGTAACGTTATTA 3193
DB |||||
3122 CCACTATGTCCTCAATGGCTTTCATGAAGCTGGCGCGCTGACAGGTAACGTAATATGA 3181
QY |||||
3194 CCACTTACCCACTCGGGGATCGGGCCACCGGGCTACGAGACTTGGCGTGGCGT 3253
DB |||||
3182 CCACTTACTCCACTCGGGGATGGGGCCACCGGGCTACGAGACTTGGCGTGGCACT 3241
QY |||||
3254 AGAGCCGCTGCTCTTCCGCCATGAGACCAAGGTCATCACTGGGGAGCAGACCCG 3313
DB |||||
3242 AGAGCCGCTGCTCTCTGACATGAGACTAACTCATCACTGGGGGCGAGACCCG 3301
QY |||||
3314 TGGGTGTGGGGAATCATCTTGGGTCTACCGCTCTCCGCCGGAAGGGGAGAGATATT 3373
DB |||||
3302 GGGGTGTGGGGAATCATCTCGGCTCTACCGTCTCCGCCGGAAGGGGAGAGATCT 3361

QY 3374 TTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTGGCGACTCTTGTGCGCCCATCACGGC 3433
DB |||||
3362 TCTAGGACCGGCGGATAGTTTGGAGAGCAGGGGTGGCGGCTCTTGTGCGCTATCACGGC 3421
QY |||||
3434 CTACTCCCAACAACCGGGGCGTACTTGGTTGATCATCTAGCTCACAAGGCCGGA 3493
DB |||||
3422 CTATTCCTCAACAACCGGGGCGTCTTGGTGTATCATCTAGCTCACAAGGTCGGGA 3481
QY |||||
3494 CAAGAACCAAGTCCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACACAAATCTTTCTCTGC 3553
DB |||||
3482 CAAGAACCAAGTCCGAAGGGAGGTTCAAGTGGTCTCTCACCGCAACGCAATCTTTCTCTGC 3541
QY |||||
3554 GACCTGCATCAACCGGCTGTCTGGAAGTCTTACCATGGCGCTGGCTCGAAGACCTTAGC 3613
DB |||||
3542 GACCTGGTCAATGGCGTGTGTGGACGCTCTACCATGGTGGCGCTCGAAGACCTTGC 3601
QY |||||
3614 CGGTCCAAAGGTCAATCACCCAAATGTACACCAATGTACCAATGTAGACAGGACCTCGTGG 3673
DB |||||
3602 CGGCCGGAAGGTCGAATCACCCAAATGTACCAATGTAGACAGGACCTCGTGGCTG 3661
QY |||||
3674 GCAGCGCCCGCGGGGCGCTCCATGCACACCATGACGCTGTGGCAGCTCGGACCTTTTA 3733
DB |||||
3662 GCCGCGCCCGCGGGGCGCTCCATGACACCGTGCACCTGCGGACGCTCGGACCTTTTA 3721
QY |||||
3734 CTTGTTCACGAGACATGCTGATGTCTTCCGCTGCGCGGAGGCGACAGCAGGGGAG 3793
DB |||||
3722 CTTGTTCACGAGGATGCTGATGTCTTCCGCTGCGCGGCGGCGGCGGACAGCAGGGGAG 3781
QY |||||
3794 TCTACTCTCCCCAGGCGGCTCTCTACTGAAAGGCTCTCGGCTGTCTCATGCTTTG 3853
DB |||||
3782 CTTGTCTTCCCGAGGCGCATCTCTACTGAAAGGCTCTCTCGGCTGGACCACTGCTTTG 3841
QY |||||
3854 CTTTCGGGGCAGCTGCGGGGCTTTCGGGCTGCTGTGTGCACCCCGGGGGTCCGCA 3913
DB |||||
3842 CTTTCGGGGCAGCTGCTGAGGATCTTCCGGGCTGCTGTGTGCACCCCGGGGGTTCGCA 3901
QY |||||
3914 GCGGTGACCTTACACCGCTTGAGTCTATGAAACTACCATGCGGCTCTCCGCTCTTAC 3973
DB |||||
3902 GCGGTGACCTTACACCGCTTGAGTCTATGAAACTACCATGCGGCTCTCCGCTCTTAC 3961
QY |||||
3974 AGACAACTCAACCCCGCGGCTGTACCGACAGCATTCGAAAGTGGCACTGTGCACGCTCC 4033
DB |||||
3962 AGACAACTCATCCCTCCGCGGTACCGCAAACTTCCAAGTGGCACATTTACAGCTCC 4021
QY |||||
4034 TACTGCGAGCGGCAAGACCAAAAGTCCGCTGCTATGCGAGCCCAAGGTAACAAGT 4093
DB |||||
4022 CACTGGCAGCGGCAAGACCAAAAGTCCGCTGCTATGCGAGCCCAAGGTAACAAGT 4081
QY |||||
4094 GCTGCTCTGAACCGCTCCGTTGCGGCCACCTTAGGGTTTGGGGGCTATATGTCCAAGGC 4153
DB |||||
4082 GCTGCTCTTAAACCGCTCCGTTGCGGCCACATTTGGGCTTTGGAGGCTATATGTCCAAGGC 4141
QY |||||
4154 ACACGGTATCGAACCCCTAAACATCAGAACTGGGGTAAGGACCAATTAACAAGGCGGCTCCAT 4213
DB |||||
4142 ACATGGCATCGAGCCTAAACATCAGAACTGGGGTAAGGACCAATCAGACCGGGCGGCCAT 4201
QY |||||
4214 TAGCTACTCCACCTATGCGAAGTCTCTTCCGCGACCGTGGCTGTTCTGGGGGCGCTATGA 4273
DB |||||
4202 CACGTACTCCACTATTTGCAAGTCTCTTCCGCGACCGTGGATGCTCCGGGGGCGCTATGA 4261
QY |||||
4274 CATCATATATGTGATGATGCTCACTCAACTGACTCGACTACCATCTTTGGGATCGGCAC 4333
DB |||||
4262 CATCATATATGTGATGATGCTCACTCAACTGACTCGACTACCATCTTTGGGATCGGCAC 4321
QY |||||
4334 AGTCTTGGACCAAGCGGAGACGGCTGGAGCGGCTCGTGTGCTCGCACCGCTTACAC 4393
DB |||||
4322 AGTCTTGGATCAGGCAGAGACGGCTGGAGCGGCTCGTGTGCTCGCACCGCACCGCC 4381
QY |||||
4394 TCCGGGATCGGTTACGCTGCGACACCCCAATATCAGGAAATAGGCTGTCTCAACATGG 4453
DB |||||
4382 TCCGGGATCGATCACCGTGCCACACCCCAACATCAGGAAAGTGGGCTGTCTCAACATGG 4441
QY |||||
4454 AGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGGCGCATCAAGGGGGGAGGCTCT 4513

Db	4442	AGAGATTCCTTCATGCGCAAGCCATCCCATTTGAGGCCATCAAGGGGGAAGGATCT	4501
Qy	4514	CATTTCCTGCAATTCGAAGAGAAATGTGACGAGCTGCGCGCAAAAGCTGACAGGCCCTCGG	4573
Db	4502	CATCTTCGCAATTCGAAGAGAAAGTGTGACGAGCTGCGCGCAAAAGCTGACAGGCCCTCGG	4561
Qy	4574	ACTGAACGCTGTAGCATATTTACCGGGCCCTTGATGTGTCGTCATACCGCCTATCGGAGA	4633
Db	4562	ACTCAATGCTGTAGCGATTTACCGGGGCTCTCGATGTGTCGTCATACCGGATAGCGAGA	4621
Qy	4634	CGTCGTTGTCGTGGCAACAGACGCTCTAATACGCGGTTTTCACCGCGGATTTGACCTCAGT	4693
Db	4622	CGTCGTTGTCGTGGCAACAGACGCTCTAATGACGCGTTTACCGCGGATTTGACCTCAGT	4681
Qy	4694	GATCGACTGCAATPACATGTGTACCCAGACAGTCGACTTCAGCTTGGATCCCACTTCAC	4753
Db	4682	GATCGACTGCAACACATGTGTACCCAGACAGTCGACTTCAGCTTGGATCCCACTTCAC	4741
Qy	4754	CATTGAGACGACACCGTGCCTCCCAAGACGCGGTGTGCGCTCGCAACCGCGAGGTAGAAC	4813
Db	4742	CATTGAGACGACAAACGTGCTCCCAAGACGCGGTGTGCGCTCGCAACCGCGAGGTAGAAC	4801
Qy	4814	TGCGAGGGTAGGATGGCATCTACAGGTTTGTGACTCCAGGAGAACCGGCCCTCGGSCAT	4873
Db	4802	TGCGAGGGTAGGATGGCATCTACAGGTTTGTGACTCCAGGAGAACCGGCCCTCGGSCAT	4861
Qy	4874	GTTTCGATTCCTTCGCTGTGTGAGTGTATGACGCGGGGTGTGCTTGGTATGAGCTCAC	4933
Db	4862	GTTTCGATTCCTTCGCTGTGTGAGTGTATGACGCGGGGTGTGCTTGGTATGAGCTCAC	4921
Qy	4934	GCCCGCTGAGACCTCGGTTAGGTTGCGGCTTACCTAAATACACAGGGTTGCGCGCTG	4993
Db	4922	GCCCGCTGAGACCTCGGTTAGGTTGCGGCTTACCTAAATACACAGGGTTGCGCGCTG	4981
Qy	4994	CGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTCACAGGGCTCACCCACATAGATGCCCA	5053
Db	4982	CGAGGACCATCTAGATTTCTGGGAGAGCGTCTTTCACAGGGCTCACCCACATAGATGCCCA	5041
Qy	5054	CTTCTCTGTCACAGACTAAACAGGCGAGAGCAACTTTCCTTACCTGCTGTCATCAAGC	5113
Db	5042	CTTCTCTGTCACAGCAACAGGCGAGAGCAACTTTCCTTACCTGCTGTCATCAAGC	5101
Qy	5114	TACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAAATGTGGAGTGTCT	5173
Db	5102	CACAGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGACCAAAATGTGGAGTGTCT	5161
Qy	5174	CATACGGCTGAACCTTACACTGCAAGGGCCAAACCCCTGCTGTATAGGCTAGGAGCGGT	5233
Db	5162	CATACGGCTGAACCTTACACTGCAAGGGCCAAACCCCTGCTGTATAGGCTAGGAGCGGT	5221
Qy	5234	CCAAAATGAGGTCACTCTCACACCCCAATAACTAAATACATCATGCGATGTCGCGC	5293
Db	5222	TCAAAATGAGGTCACTCTCACACCCCAATAACTAAATACATCATGCGATGTCGCGC	5281
Qy	5294	TGACCTGGAGGTCGTCACTAGCACCTGGGTGCTGGTAGGGGGAGTCCCTTGAGCTTTGGC	5353
Db	5282	TGACCTGGAGGTCGTCACTAGCACCTGGGTGCTGGTAGGGGGAGTCCCTTGAGCTTTGGC	5341
Qy	5354	CGCATATGCTGTACGACAGGCACTGTGTCATTTGTGGGAGGATCATCTTTGTCGGGAA	5413
Db	5342	CGCGTATGCTGTACGACAGGCACTGTGTCATTTGTGGGAGGATCATCTTTGTCGGGAG	5401
Qy	5414	GCCAGCTGTGTTCCCGACAGGGAAGTCTCTACAGGAGTTCGATGAGTGAAGAGTG	5473
Db	5402	GCCAGCTGTGTTATCCCGACAGGGAAGTCTCTCTACAGGAGTTCGATGAGTGAAGAGTG	5461
Qy	5474	TGCTCTCAAACTCTCTTACATCGAGCAGGGAATGACGCTCGCGAGCAATTCGAAGCAAAA	5533
Db	5462	TGCTCTCAAACTCTCTTACATCGAGCAGGGAATGACGCTCGCGAGCAATTCGAAGCAAAA	5521
Qy	5534	GGCGCTCGGGTTGTGAAACGGGCCAACGAACGAGGAGGCTGCTCTCCCGTGGTGA	5593

Db	5522	GGCGCTCGGATTGCTGCAAAACAGCCACCAAGCAAGCGGAGGCTGCTCTCCCGTGGTGA	5581
Qy	5594	GTCCAAAGTGGCGAGCCCTTTGAGACCTTTCTGGGGGAAGCACATGTGGAATTTCAATCAGCGG	5653
Db	5582	GTCCAAAGTGGCGAGCCCTTTGAGGTTCTTTGGGGGAACACATGTGGAATTTCAATCAGCGG	5641
Qy	5654	AATACAGTACTAGCAGGCTTATCCACTCTGCTGGAACCCCGCGATAGCATCATTGAT	5713
Db	5642	GATACAGTACTTGGCAGGCTTATCCACTCTGCTTGGAAACCCCGCGATAGCATCATTGAT	5701
Qy	5714	GGCAITTTACAGCTTCTATCATTAGCCCGCTCACACCCCAAAACACCCCTCTCTGTTTAAAT	5773
Db	5702	GGCTTTTACAGCTTCTATCACCAGCCGCTCACACCCCAAAATACCCCTCTCTGTTTAAAT	5761
Qy	5774	CTTGGGGGATGCGTGGCTGCCCAACTCGTCTCTCCAGCGCTGCGTACGCTTTCTGTTGGG	5833
Db	5762	CTTGGGGGATGCGTGGCTGCCCAACTCGTCTCTCCAGCGCTGCGTACGCTTTCTGTTGGG	5821
Qy	5834	CGCCGCATCGCCGAGCGGCTGTTGGCAGCATAGGCTTTGGGAAGGTGCTCGTGGACAT	5893
Db	5822	CGCCGCATTTGCCGCTGCGGCCGTTTGGCAGCATAGGCTTGGGAAGGTACTTGTGGACAT	5881
Qy	5894	CTTGGCGGGTATGCGGCGAGGGGTAGCCGCGCACTCGTGGCCCTTTAAAGGTCTATGAGCGG	5953
Db	5882	TCCTGGCGGCTATGCGGCGGGGTGCTGGCGCACTCGTGGCCCTTTAAAGGTCTATGAGCGG	5941
Qy	5954	CGAGGTGCCCTCCACCGAGACCTGTGCAACTTACTCTCCCTGCCATCTCTCTCTGTGTGC	6013
Db	5942	CGAGATGCCCTCCACTGAGGATCTGGTTAATTTACTTCCCTGCCATCTCTCTCTGTGTGC	6001
Qy	6014	CTTGGTCTCGGGTGTGTCGACGCAATCTCGCTCGGACGTCGCGCGGAGGG	6073
Db	6002	CTTGGTGTCTCGGGTGTGTCGACGCAATCTCGCTCGGACGTCGCGCGGAGGG	6061
Qy	6074	GGCTGTGACGTGATGAACCGGCTGATAGCGTTTCGCTTCGCGGGGTAAACACGCTCTCCCC	6133
Db	6062	GGCTGTGACGTGATGAACCGGCTGATAGCGTTTCGCTTCGCGGGGTAAACACGCTCTCCCC	6121
Qy	6134	TACGCACTATGCTGCTGAGAGCGCTGACGACGCTGTCACGCTGCTCTCTCTAGCCT	6193
Db	6122	CACGCACTATGCTGCTGAGAGCGCTGACGACGCTGTCACGCTGCTCTCTCTAGCCT	6181
Qy	6194	TACATCACTCAACTGCTGAGCGGCTCACACGAGTGAATTAATGAGACGCTCTCTAGCCT	6253
Db	6182	TACATCACTCAACTGCTGAGAGGCTTCTCATGTTGAATTAATGAGACGCTCTCTAGCCT	6241
Qy	6254	ATGCTCGGCTCGTGGCTAAGGATGTTTGGGATTTGGATATGACGCTGTTGACTGACTT	6313
Db	6242	TTGTTTCGGCTCGTGGCTAAGGATGTTTGGGATTTGGATATGACGCTGTTGACTGACTT	6301
Qy	6314	CAAGACCTGCTCCAGTCCAAACTCTCTGCGCGGTTACCGGAGTCCCTTTCTGTGTATG	6373
Db	6302	CAAGACTTGGCTCCAGTCCAAACTCTCTGCGCGGTTACCGGAGTCCCTTTCTGTGTATG	6361
Qy	6374	CCAAACGGGTTAAGGAGTCTGGCGGGGGAGCGGATCATGCAAAACACCTGCCCATG	6433
Db	6362	CCAAACGGGTTAAGGAGTCTGGCGGGGGAGCGGATCATGCAAAACACCTGCCCATG	6421
Qy	6434	CGGAGCACATGCCCGGACATGTCAAAAACGGTTCCATGAGGATCTGAGGCTCTAGAC	6493
Db	6422	TGGAGCACATCACCGGACATGTCAAAAATGGCTCCATGAGGATTTGTGGGCGCAAAAC	6481
Qy	6494	CTGCAGCAACACGTTGGCAGCAAGTTTCCCATCAACGACATACACACGGGACCTTGAC	6553
Db	6482	CTGCAGCAACACGTTGGCAGTGGAAATTTCCCATCAACGACATACACACGGGCGCTGAC	6541
Qy	6554	ACCTCTCCCGCGCCAACTATTTCAGGGGCTATGCGGGTGGCTGCTGAGGAGTACGT	6613
Db	6542	GCCCTCCCGCGCGGAACTATTTCAGGGGCTATGCGGGTGGCTGCTGAGGAGTACGT	6601
Qy	6614	GGAGGTTACCGGTTGGGGATTTCCACTAGTGACCGGACATGACCACTGACCACTGATAA	6673
Db	6602	GGAGGTTACCGGTTGGGGATTTCCACTAGTGACCGGACATGACCACTGACCACTGATAA	6661

QY 6674 GTGCCATGCCAGGTTCCGGCCCCCGAAATCTTTACCGAGGTGGATGAGTGCCTGTTGCA 6733
DB |||||
6662 ATGCCCATGCCAGTTTCAGCCCCCTGAATTTTTCAGCGAGGTGATGAGTACGGTTGCA 6721
QY 6734 CAGGTACGCTCCGGCGTGCAAACTCTTCTACGGGAGGACGTCACGTTTCCAGGTGCGGCT 6793
DB |||||
6722 CAGGTATGCTCCAGTGTGCAAACTCTCTACGAGAGGAGTCTGTTTCCAGGTGCGGCT 6781
QY 6794 CAACCAATACTTGGTTCGGGTCCGAGTCCCATGCGAGCCCGAACCAGGACTAAAGTGCCT 6853
DB |||||
6782 CAACCAAGTACCTGGTTCGGGTCCAGCTCCCATGTGTAGCCCCGAACCGGATGTGGCAGTGCCT 6841
QY 6854 TACTTCCATGTCTACCGATCCCTCCCAATTTACAGCAGACGGCTAAGCGGTAGGCTGGC 6913
DB |||||
6842 CACTTCCATGTCTACCGACCCCTCTCATATTACAGCAGACGGCCAGCGGTAGGCTGGC 6901
QY 6914 TAGAGGGTCTCCCCCTCTTTAGCCAGCTCATAGCTAGCCAGTTGTCTGCGCCCTCTTT 6973
DB |||||
6902 CAGGGGTCTCCCCCTCTTTGGCCAGCTCTTCAGCTAGCCAGTTGTCTGCGCCCTCTTT 6961
QY 6974 GAAGCGACATGCACTACCCACCATGACTCCCGGACCGTGCACCTCATCGAGGCCAACCT 7033
DB |||||
6962 GAAGCGACATGCACTACCCACCATGACTCCCGGACCGTGCACCTCATCGAGGCCAACCT 7021
QY 7034 CTTGTGGCGGAGGAGATGGGCGGAAACATCACTCGGTGGAGTCAGAGAAATAAGGTAGT 7093
DB |||||
7022 CTTGTGGCGGAGGAGATGGGCGGAAACATCACTCGGTGGAGTCAGAGAAATAAGGTAGT 7081
QY 7094 AATCTGGAATCTTTGCAACCGCTTTCAGCGGAGGGGATGAGAGGAGATATCGTGC 7153
DB |||||
7082 AATCTGGACTCTTTTCGATCCGATTCGGGCGGTGGAGATGAGAGGAAATATCGTCCC 7141
QY 7154 GCGGAGATCTTCGGAATAATCAGAAAGTTCCCTCAGGTTGGCCCATATGGGACGCCC 7213
DB |||||
7142 GCGGAGATCTTCGGAATAATCAGAAAGTTTCCTCCGACGCTTGGCCCATATGGGACGCCC 7201
QY 7214 GGACTACAACTCTCCACTGCTAGAGTCTCTGGAAGGACCGGACTACGTCCTTCGGTGGT 7273
DB |||||
7202 GGATTAACAACCTCTCACTGCTAGAGTCTCTGGAAGGACCGGACTACGTCCTCCCGTGGT 7261
QY 7274 ACACGGATGCCCATTTGCCACTTACCAAGGCTCTCTCAATACCACTTCCACGGAGAAAG 7333
DB |||||
7262 ACACGGGTGCCCTTTGGCATCTTACCAAGGCCCTCCCAATACCCTCCACGGAGGAAG 7321
QY 7334 GACGTTGTCTGACAGAAATCCAAATGTGTCTCTGCTTGGCGGAGCTCGCCACTAAGAC 7393
DB |||||
7322 GACGTTGTCTGACAGAGTCCACCGTGTCTTCTGCTTGGCGGAGCTCGCTACTAAGAC 7381
QY 7394 CTTTCGTAGCTCCGGATCGTCCGCGTGTGATAGCGGACCGGCTCTTCTGACCT 7453
DB |||||
7382 CTTTCGACCTCCGGTCTCGGCGTGTGACAGCGGACCGGCTCTTTCGATCA 7441
QY 7454 GGCCTCCGACGAGCGGTGACAAAGGATCCGACGTTGAGTCTGCTCTCATGCCCTCCCT 7513
DB |||||
7442 GGCCTCCGACGAGCGGACAAAGGATCCGACGTTGAGTCTGCTCTCATGCCCTCCCT 7501
QY 7514 TGAAGGGAGCGCGGGGAACCCGATCTCAGCGACGGGCTCTTGGTCTACCGTGAAGTGA 7573
DB |||||
7502 CGAGGGAGAGCGAGGGGAACCCGACCTCAGCGACGGGCTCTTGGTCTTACCGTGAAGGGGA 7561
QY 7574 GGCTAGTGAAGGATGCTGCTGCTGCTCAATGTCTTATAGTCGACAGGCGCCCTGATCAC 7633
DB |||||
7562 AGCTGTGAGGACGCTGCTGCTGCTCAATGTCTTATATAGGACAGGTCCTTGTATCAC 7621
QY 7634 GGCATGCGCTGGGAGGAAAGTAAAGTGCCTCATCAACCGGTTGAGCAACTCTTTGCTGG 7693
DB |||||
7622 GGCATGCGCTGGGAGGAGCAAGTTGGCCATCAATCCGTTGAGCAACTCTTTGCTGG 7681
QY 7694 TCACCAACAATGGTCTTACGCCAACAATCCCGGACGCGCAAGCTCCGGCAGAAAGGT 7753
DB |||||
7682 TCACCAACAATGGTCTTACCCAACAATCTCGCAGCGCAAGTCTTCGCGCAGAAAGGT 7741

QY 7754 CACCTTTGACAGATTGCAAGTCTCGATGATCATTTACCGGACGTTACTCAAGGAGATGAA 7813
DB |||||
7742 CACCTTTGACAGACTGCAAGTCTCGACGACCACTACCGGACGTTCTCAAGGAGATGAA 7801
QY 7814 GGCAGAGGCTCCACAGTTAAGGCTAAGCTTCTATCTATAGGAGGCTCTCGACTGAC 7873
DB |||||
7802 GGCAGAGGCTCCACAGTTAAGGCTAAGCTTCTATCTATAGGAGGCTCTCGAACTGAC 7861
QY 7874 GCGCCCATCTCGGCCAAATCCAAATTTGGCTATGGGCAAAAGACGTCGGGAACCTATC 7933
DB |||||
7862 GCGCCCATCTCGGCCAAATCCAAATTTGGCTA CGGGCGAAGGACGTCGGAGCTATC 7921
QY 7934 CAGCAGGCGCTTAAACCAATCCGCTCCGCTGGGAGACTTGTGGAAGACACTGGAAC 7993
DB |||||
7922 CAGCAGGCGCTTAAACCAATCCGCTCCGCTGGGAGACTTGTGGAAGACACTGGAAC 7981
QY 7994 ACCAATTTGACACCACTCATGCGCAAAAGTGAAGTTTCTGCGTCCAAACGAGAAAGG 8053
DB |||||
7982 ACCAATTTGATACCACTCATGCGCAAAAGTGAAGTTTCTGCGTCCAAACGAGAAAGG 8041
QY 8054 AGCGCGAAGCAGCTCGCTTATCGTATTCAGACCTGGAGTTCGTGTATGCGAGAA 8113
DB |||||
8042 AGCGCGAAGCAGCTCGCTTATCGTATTCAGACCTGGGCTGCGTGTATGCGAGAA 8101
QY 8114 GATGCGCTTTACGACGCTGCTCCACCTTCTCAGCGCTGATGGCTCCTCATACGG 8173
DB |||||
8102 GATGCGCTTTACGACGCTGCTCCACCTTCTCAGCGCTGATGGGCTCCTCATACGG 8161
QY 8174 ATTTCAATACTCCCCAAGCAGCGGCTCAGTTCCTGTGTAATACCTGGAATCAAAGAA 8233
DB |||||
8162 ATTCAGTACTCTCTGCGGACGCGGCTCAGTTCCTGTGTAATACCTGGAATCAAAGAA 8221
QY 8234 ATGCGCTATGGCTTCTCATATGACCGCTGTTTTCAGCTCAACGGTCACTGAGAGTGA 8293
DB |||||
8222 ATGCGCTATGGCTTCTCATATGACCGCTGTTTTCAGCTCAACGGTCACTGAGAAATGA 8281
QY 8294 CATTCGTTGAGGAGTCAATTTACCAATGTGTGACTTGGCCCCGAGCGGCAGACAGCG 8353
DB |||||
8282 CATCGTACTGAGGAATCAATTTACCAATGTGTGACTTGGCCCCGAGCGGCAGCG 8341
QY 8354 CATAGGTCGCTCAGAGCGGCTTTTACATCGGGGCTCCCTGACTAATCAAAAGGGA 8413
DB |||||
8342 CATAGGTCGCTCAGAGCGGCTTTTATGTGCGGGCTCCCTGACTAATTCGAAAGGGGA 8401
QY 8414 GAACTGCGGTTATCGCGGTCGCGCAAGTGGCTGTGACGCTAGCTGCGGTAAATAC 8473
DB |||||
8402 GAACTGCGGTTATCGCGGTCGCGCAAGTGGCTGTGACGCTAGCTGCGGCAACAC 8461
QY 8474 COTCAATGTTACTTTGAAGGCCACTGCGACCTGTGAGCTGCAAGCTCCAGGACTGCAC 8533
DB |||||
8462 COTCAATGTTACTTTGAAGGCCACTGCGGCTGTGAGCTGCAAGCTCCAGGACTGCAC 8521
QY 8534 GATGTCGTGAACGAGAGACGCTTGTGTTATCTGTGAAAGCGCGGGAACCCAGGAGA 8593
DB |||||
8522 GATGTCGTGAACGAGAGACGCTTGTGTTATCTGTGAGAGTGGCGGAACCCAGGAGA 8581
QY 8594 TCGCGGCGCTTACGAGCTTACGAGGCTATGACTAGTATTCGCGCCCCCGGGGA 8653
DB |||||
8582 TCGCGGCGCTTACGAGCTTACGAGGCTTACGAGGCTATGACTAGTATTCGCGCCCCCGGGGA 8641
QY 8654 TCGCGCCCCAACAGAAATACGACTCGGAGCTGATAAATCATGTTCTCTCAATGTGTGCT 8713
DB |||||
8642 CCGCGCCCCAACCAGAAATACGACTTGGAGCTGATAACGTCATGCTCTCTCAATGTGTGCT 8701
QY 8714 CCGCGCAATGATCTGGCAAAAGGCTATACCTCAGCCGCTGACCCCAACCCCT 8773
DB |||||
8702 CCGCGCAATGATCTGGCAAAAGGCTATACCTCAGCCGCTGACCCCAACCCCT 8761
QY 8774 TGCAGGCTGCTGGGAGACAGCTAGACACACTCCAATCAACTCTTGTCTAGGCAATAT 8833
DB |||||
8762 CCGAGCGGCTGCTGGGAGACAGCTAGACACACTCCAATCAACTCTTGTCTAGGCAATAT 8821
QY 8834 CATCATGTATGCGCCCCCCTTATGGGCAAGGATGATTTGTGATGACTCACTTTTCTCCAT 8893

DB 8822 CATCATGATATGCGCCACCTATGGCGGAGGATGATCTGATGACTCATTTCTTCTCTAT 8881
QY 8894 CCTTCTAGCTCAAGAGCACTTCAAAAAGCCCTGGATTGTGAGATCTACGGGGCTTGCTA 8953
DB 8882 CCTTCTAGCTCAGAGCACTTGAAGAAGCCCTGGATTGTGAGATCTACGGGGCTTGTTA 8941
QY 8954 CTCATGAGCCACTTGACCTACCTCAGATCATTTGAAGCATCCATGGTCTTAGCCGATT 9013
DB 8942 CTCATTTAGCCACTTGACCTACCTCAGATCATTTGAAGCATCCATGGTCTTAGCCGATT 9001
QY 9014 TACATCCACAGTACTCTCCAGTGAGATCAATAGGTGGTTCATGCTCAGGAACT 9073
DB 9002 TTCACCTCCAGTACTCTCCAGTGAGATCAATAGGTGGTTCATGCTCAGGAACT 9061
QY 9074 TGGGTACCACTCTGGAACCTTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACT 9133
DB 9062 TGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACT 9121
QY 9134 GTCCAGGGGGAGGCGCCCACTTTGTGCGAGATACCTCTTTAACTGGGCAGTAAGGAC 9193
DB 9122 GTCCAGGGGGAGGCGCTGCCACTTGGCGCAAGTACTCTTCAACTGGGCAGTAAGGAC 9181
QY 9194 CAAGCTTAACTCACTCTCAATCCGCGCGGTCCAGCTGACTTGTCTGGCTGCTCGT 9253
DB 9182 CAAGCTTAACTCACTCTCAATCCGCGCGGTCCAGCTGACTTGTCTGGCTGCTCGT 9241
QY 9254 CGCTGGTTACAGCGGGGAGACATATATCAGCTGTCTGCGCGACCCCGCTGGTT 9313
DB 9242 TGTGGTTTACAGCGGGGAGACATATATCAGCTGTCTGCGCGACCCCGCTGGTT 9301
QY 9314 TCCGTTGTGCTACTCTCTACTTTCTGTAGGGTAGGCAITTTACCTGCTCCCCCAACCGATG 9373
DB 9302 CATGTTGTGCTACTCTCTACTTTCTGTAGGGTAGGCAITTTACCTGCTCCCCCAACCGATG 9361
QY 9374 AACGGGAGCTAACCACTCCAGGCTTTAAGCCATTTCTGTTTCTGTTTCTGTTT 9424
DB 9362 AACGGGAGCTAACCACTCCAGGCAATAGGCCATTCCTCTTTTCTGTTTCTGTTT 9412

RESULT 10
AAD49655
ID AAD49655 standard; DNA; 9413 BP.
XX AC AAD49655;
XX DT 24-MAR-2003 (first entry)
XX DE Hepatitis C virus (HCV) DNA.
XX KW Amyloidosis; haemophilia; Alzheimer's disease; atherosclerosis; cancer;
KW gigantism; dwarfism; hypothyroidism; hyperthyroidism; cystic fibrosis;
KW autoimmune disorder; aging; inflammation; diabetes; obesity; anorectic;
KW neurodegenerative disorder; Parkinson's disease; gene therapy; virucide;
KW haemostatic; antibacterial; nootropic; neuroprotective; cytostatic;
KW fungicide; Hepatitis C virus; HCV; gene; ds.
XX OS Hepatitis C virus.
XX PN WO200283953-A1.
XX PD 24-OCT-2002.
XX PF 11-APR-2002; 2002WO-US011757.
XX PR 11-APR-2001; 2001US-0282965P.
XX PA (PTCT-) PTC THERAPEUTICS INC.
XX PI Rando R, Welch E;
XX DR WPI; 2003-075561/07.

PT Identifying a test compound that binds to a target RNA molecule for
PT treating or preventing amyloidosis, hemophilia, cancer, gigantism,
PT diabetes, by contacting a detectably labeled target RNA molecule with a
PT library of test compounds.
PS Example; Page 64-68; 152pp; English.
XX The invention relates to a method for identifying a test compound that
XX binds to a target RNA molecule, which comprises contacting a detectably
XX labelled target RNA molecule with a library of test compounds under
XX conditions that permit direct binding of the labelled target RNA to a
XX member of the library of test compounds so that a detectably labeled
XX target RNA-test compound complex is formed. The method is useful for
XX screening libraries of compounds for those that are selectively bind to a
XX pre-selected target RNA. The compounds are useful for inhibiting the
XX formation of a specific bound RNA: host cell factor complexes in vivo.
XX They are also useful for treating or preventing diseases associated with
XX overproduction or decreased protein function, such as amyloidosis,
XX haemophilia, Alzheimer's disease, atherosclerosis, cancer, gigantism,
XX dwarfism, hypothyroidism, hyperthyroidism, autoimmune disorders, aging,
XX inflammation, cystic fibrosis, diabetes, obesity, neurodegenerative
XX disorders, Parkinson's disease or infections (bacterial, viral, fungal).
XX The invention is also used in gene therapy. The present sequence is
XX Hepatitis C virus (HCV) DNA. This sequence is used to illustrate the
XX method of the invention.
SQ Sequence 9413 BP; 1886 A; 2816 C; 2694 G; 2017 T; 0 U; 0 Other;
Query Match 86.6%; Score 8311.8; DB 8; Length 9413;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 8724; Conservative 0; Mismatches 687; Indels 0; Gaps 0;
QY 14 TGGGGCGGACATCCACCATGAATCACTCCCTGTGAGGAATCTGTCTTACGGAGAA 73
DB 2 TGGGGCGGACATCCACCATGAATCACTCCCTGTGAGGAATCTGTCTTACGGAGAA 61
QY 74 AGGCTTAGCCATGGCGTTAGTATGATGTCGTGACGCTCCAGGACCCCTCCCGGG 133
DB 62 AGGCTTAGCCATGGCGTTAGTATGATGTCGTGACGCTCCAGGACCCCTCCCGGG 121
QY 134 AGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGACGAGACCGGTGTC 193
DB 122 AGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGACGAGACCGGTGTC 181
QY 194 TTTCTTGGATCAACCCCGCTCAATGCTGAGATTTGGCGGTGCCCGGAGACTGTAG 253
DB 182 TTTCTTGGATCAACCCCGCTCAATGCTGAGATTTGGCGGTGCCCGGAGACTGTAG 241
QY 254 CCGAGTAGTGTGGGTGCGAAAGGCGCTTGTGTTACTGCTGATAGGTTGCTTCGGAGTG 313
DB 242 CCGAGTAGTGTGGGTGCGAAAGGCGCTTGTGTTACTGCTGATAGGTTGCTTCGGAGTG 301
QY 314 CCGCGGAGGTCTGTTAGACCGGTGACCATGACGACCAATCTTAACTCAAGAAAC 373
DB 302 CCGCGGAGGTCTGTTAGACCGGTGACCATGACGACCAATCTTAACTCAAGAAAC 361
QY 374 CAACAGTAACACCAACCGCGCCACAGGAGCTCAAGTTCCCGGGCGGTGTCAGATCGT 433
DB 362 CAACAGTAACACCAACCGCGCCACAGGAGCTTAAAGTTCCCGGGCGGTGTCAGATCGT 421
QY 434 TGGTGGAGTTTACCTGTTGCGCGCAGGGGCCCCAGGTTGGGTGTCGCGGACTAGGAA 493
DB 422 TGGTGGAGTTTACCTGTTGCGCGCAGGGGCCCCAGGTTGGGTGTCGCGGACTAGGAA 481
QY 494 GGCTTCCGAGCGGTGCGAAACCTCTGGAAGCGGACCACTATCCCAAGGTCGCGGACC 553
DB 482 GACTTCCGAGCGGTGCGAAACCTCTGGAAGCGGACCACTATCCCAAGGTCGCGGACC 541
QY 554 CGAGGCGAGGCTGGGCTCAGCCCGGTGATCCCTTGGCCCTCTATGCAATAGGGGCT 613
DB 542 CGAGGCGAGGCTGGGCTCAGCCCGGTGATCCCTTGGCCCTCTATGCAATAGGGGCT 601
QY 614 GGGGTGGGCGAGGATGGCTCTCTGTACACCCCGGGCTCCCGGCTAGTTGGGGCCCCACGGA 673

QY	2834	GTCAACATACATAAAGTGTCTTCTACTAGGCTCATATGTTGGTTACAATACCTTTATCAC	2893
DB	2822	GTCAACATACATAAAGTGTCTCTCGTAGGCTCATATGTTGGTTACAATATTTATCAC	2881
QY	2894	CAGAGCGAGCGCACATCAAGTGTGGTCCCGCCCTCAAGTTCGGGGAGCGCGA	2953
DB	2882	CAGAGCGAGCGCACATTCAAGTGTGGTCCCGCCCTCAAGTTCGGGGAGCGCGA	2941
QY	2954	TGCCATCATCTCTCAAGTGTGGGTTCAATCAGAGTTAAATTTTGAATCAACCAAACT	3013
DB	2942	TGCCATCATCTCTTATACATGCGGGTCAATCAGAGTTAAATTTTGAATCAACCAAACT	3001
QY	3014	CTGCTCGCCATACCTCGGCCGCTCATGTTGCTCCAGGCTGGCATAAACGAGATGCGGTA	3073
DB	3002	CTGCTCGCCATACCTCGGCTCGGCTCATGTTGCTCCAGGCTGGCATAACTAGATGCGGTA	3061
QY	3074	CTTTCGCGGCTCAAGGGCTCATTCGTGATGATGTTAGTGGAAAGTCCCGCGGG	3133
DB	3062	CTTTCGCGGCTCAGGGGCTCATTCGTGATGATGTTAGTGGAAAGTCCCGTGGAG	3121
QY	3134	TCATTATGTCAAATGCTTTCTGAAGCTGGGCGCTGACAGGTACGTACGTTTATAA	3193
DB	3122	CCACTATGTCAAATGCGCTTCATGAAGCTGGGCGCTGACAGGTACGTACGTTATGA	3181
QY	3194	CCATCTTACCCACTCGCGGACTGGGCCACCGGGGCTTACGAGACCTTCGGTGGCGGT	3253
DB	3182	CCATCTTACTCCACTCGCGGATGGGCCACCGGGGCTTACGAGACCTTCGGTGGCAGT	3241
QY	3254	AGAGCCGCTGCTTCTTCGCGCATGGAGACCAAGGTATCATCCTGGGGAGGACACCGC	3313
DB	3242	AGAGCCGCTGCTTCTCTGACATGGAGACTTAAACTCATCACCTGGGGGCGAGACACCGC	3301
QY	3314	TGCGTGTGGGACATCATCTGGGTCTACCGTCTCCGCGCGAGGGGGAAGAGATATT	3373
DB	3302	GGCGTGTGGGACATCATCTGGGTCTACCACTCTCCGCGCGAAGGGGGAAGAGATATT	3361
QY	3374	TTTCGGGACCGCTCATAGTCTCGAAGGCAAGGCTGGCGACTCTCTTCGCGCCATCACCGC	3433
DB	3362	TCTAGGACCGGCCGATAGTTTGGAGAGCAGGGGTGGGGCTCTCTTGGCGCTATCACGC	3421
QY	3434	CTACTCCCAACAAACGCGGGCGTACTTGGTTGATCATATCATAGCTCTACAGCGCGGGA	3493
DB	3422	CTATTCCCAACAAACGCGGGCGCTTCTGGGTGATCATCATAGCTCTACAGGTCTCGGA	3481
QY	3494	CAAGAACAGGTGCAAGGGGAGTTCAAGTGGTTCTACCGCAACAACTCTTCTCTGGC	3553
DB	3482	CAAGAACAGGTGCAATGGGGAGGTTCAAGTGGTCTCTCCACCGCAACGCAATCTTCTCTGGC	3541
QY	3554	GACCTGCATCAACGCGGTGTCTGACTGTCTACCATGGCGCTGGCTCGAAGACCTTAGC	3613
DB	3542	GACCTGCCTCAATGGCGTGTGTGGACCGTCTTACATGGTGGCGCTCGAAGACCTTAGC	3601
QY	3614	CGGTCCAAAGGTCCAACTCACCAAAATGACCAATGATAGACCTTGGACCTCGCGGTG	3673
DB	3602	CGGCCCGGAAGGTCCAACTCACCAAAATGATACCAATGTAGACCAAGGACCTCGTGGGTG	3661
QY	3674	CGAGCGGCCCGCGGGCGGCTCATGACACCATGAGCTGTGGACCTCGGACCTTTA	3733
DB	3662	GCCGCGGCCCGCGGGCGGCTCATGACACCATGAGCTGTGGACCTCGGACCTTTA	3721
QY	3734	CTTTCGTCAGGACATGTGATGTCATTCGCGTGGCGCGGAGCGGACAGCAGGGGAAG	3793
DB	3722	CTTTCGTCAGGACATGTGATGTCATTCGCGTGGCGCGGAGCGGACAGCAGGGGAAG	3781
QY	3794	TCTACTCTCCCGCGCGGCTCTCTACTCTGAAGGCTCTCGGGTGGTCCATTCCTTTG	3853
DB	3782	CTGCTTTTCCCGCGGCGGCTCTCTACTCTGAAGGCTCTCTCGGGTGGTCCATTCCTTTG	3841
QY	3854	CCCTTCGGGGACGCTGTGGGCGTCTTTCGGGCTGTGTGTGCAACCGGGGGGTGCGGAA	3913
DB	3842	CCCTTCGGGGACGTTGTAGGCACTTTCGGGCTGTGTGTGCAACCGCGGGGGTGTGCA	3901
QY	3914	GGCGTGGACTTTCATACCGGTTGACTATGAAACTACCATGCGGTCTCCGGTCTTTCAC	3973
DB	3902	GGCGTGGACTTTCATACCGGTTGACTATGAAACTACCATGCGGTCTCCGGTCTTTCAC	3961
QY	3974	AGACAACTCAACCCCGCGGCTGTACCGCAGACATTTCAAAGTGGCACATCTGCACGCTCC	4033
DB	3962	AGACAACTCATCCCTCGGCGGTATACCGCAAACTTTCAAAGTGGCACATTTACAAGCTCC	4021
QY	4034	TAATGGCAGCGGCAAGAGCACCAAAAGTCCGGCTGGTATGAGCCCAAGGGTACAAAGGT	4093
DB	4022	CATGGCAGCGGCAAGAGCACCAAAAGTCCGGCTGCATATGAGCCCAAGGGTACAAAGGT	4081
QY	4094	GCTCGTCTGAAACCGTTCGCGCCACTTTAGGGTTTGGGGCTATATGTCCAAGGC	4153
DB	4082	GCTCGTCTGAAACCGGCTCGGTTCCGCCACATTTGGGCTTTGGAGCGTATATGTCCAAGGC	4141
QY	4154	ACACGATATCGACCCCTAAACATCAGAACTCGGGTAAGGACCATTAACACCGGGGCTCCAT	4213
DB	4142	ACATGGCATCGAGCCTTAACTCAGAACTGGGGTAAGGACCATTAACACCGGGGCTCCAT	4201
QY	4214	TAGCTACTCACTATGGCAAGTTCTTCCGCGACGCTGGCTTTCTGGGGGCGCTTATGA	4273
DB	4202	CAGCTACTCACTATTTGCAAGTTCTTCCGCGACGCTGGATGCTCCGGGGGCGCTTATGA	4261
QY	4274	CATCATAATATGTAGTGGCCTCAACTGACTCGACTTACCCTCTTGGGCTACGGCAC	4333
DB	4262	CATCATAATATGTAGTGGCCTCAACTGACTCGACTTACCCTCTTGGGCTACGGCAC	4321
QY	4334	AGTCTCGGACCAAGCGGAGACGCGCTGGAGCGGCTCGTGTGCTCGCCACCGCTACACC	4393
DB	4322	AGTCTCGGATCAGGACAGACGCGCTGGAGCGGCTCGTGTGCTCGCCACCGCTACACC	4381
QY	4394	TCGGGATCGGTTACCGGTGCCACACCCCAATATTCAGAGGAATAGGCTGTCCAAACATGG	4453
DB	4382	TCGGGATCGATCACCGTGCCACACCCCAATATTCAGAGGAATAGGCTGTCCAAACATGG	4441
QY	4454	AGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGGCCATCAAGGGGGGAGGCACTCT	4513
DB	4442	AGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGGCCATCAAGGGGGGAGGCACTCT	4501
QY	4514	CATTTCTGCTCATTTCCAAAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG	4573
DB	4502	CATCTTCTGCCATTTCCAAAGAAAGTGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG	4561
QY	4574	ACTGAACGCTGTAGCATATTTACCGGGGCTTGTATGTCTCGTCTATACCGCTATCGGAGA	4633
DB	4562	ACTCAATGCTGTAGCGTATTTACCGGGCTCTCGATGTCTCGTCTATACCGCTATCGGAGA	4621
QY	4634	CGTGGTGTGTGGCAACACACGCTCTAATGACGGGTTTCCCGGCGATTTTGGACTCAGT	4693
DB	4622	CGTGGTGTGTGGCAACACACGCTCTAATGACGGGTTTACC CGGGGACTTTTGGACTCAGT	4681
QY	4694	GATCGACTGCAATATGTTGTCTCCACAGACAGTCTCAGCTTTGGATCCCACTTTCAC	4753
DB	4682	GATCGACTGCAACACATGTGTCTCCACAGACAGTCTCAGCTTTGGATCCCACTTTCAC	4741
QY	4754	CATTGAGACGACGACCGTGGCCCAAGACGCGGTGTCCGCTGTGCAACCGGAGGTAGAAC	4813
DB	4742	CATTGAGACGACGACGCTGGCCCAAGACGCGGTGTCCGCTGTGCAACCGGAGGTAGAAC	4801
QY	4814	TGGCAGGGGTAGGATGGCATCTACAGTTTGTGACTCCAGGAGAACCGGCCCTCGGSCAT	4873
DB	4802	TGGCAGGGGTAGGATGGCATCTACAGTTTGTGACTCCAGGAGAACCGGCCCTCGGSCAT	4861
QY	4874	GTTTCGATCTTCGGTCTGTGTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT	4933
DB	4862	GTTTCGATCTTCGGTCTGTGTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT	4921
QY	4934	GCCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACTTAAATACACAGGGTTGCCGCTGTG	4993
DB	4922	GCCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACTTAAATACACAGGGTTGCCGCTGTG	4981
QY	4994	CCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTCACAGGCGCTCACCCACATAGATGCCCA	5053

||||| 4982 CCAGGACCACCTAGAGTTCTGGGAGAGGCTTTCAAGGCCTCACCCACATAGATGCCCA 5041
QY ||||| 5054 CTTTCTGTCCACAGCTAAACAGCGAGGACAACTTCTTACCTGGTGGCATACAAGC 5113
Db ||||| 5042 CTTTCTGTCCACAGCTAAACAGCGAGGAGAACCTCCCTACCTACCTGGTAGCATACAAGC 5101
QY ||||| 5114 TACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAAATGTGGAAGTGCT 5173
Db ||||| 5102 CACAGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGACCAAAATGTGGAAGTGCT 5161
QY ||||| 5174 CATACGGCTGAACCTTACATCGACGGGCCAACACCCCTGTGTATAGCTAGGAGCCGT 5233
Db ||||| 5162 CATACGGCTTAAAGCCACACTGCATGGGCCAACCGCCCTGTGTACAGGCTAGGAGCCGT 5221
QY ||||| 5234 CCAAAATGAGGTCCTCTCACACACCCCACTAACTAAATACATCATGATGGATGTCGC 5293
Db ||||| 5222 TCAAAATGAGGTCCTCTCACACACCCCACTAACTAAATACATCATGATGGATGTCGC 5281
QY ||||| 5294 TGACCTGGAGTCTGCTACCTAGCACCTGGGTGCTGTAGGCGAGTCTTTCAGCTTTTGGC 5353
Db ||||| 5282 TGACCTGGAGTCTGCTACCTAGCACCTGGGTGCTGTAGGCGAGTCTTTCAGCTTTTGGC 5341
QY ||||| 5354 CGCATCTGCTGACGACAGGAGTGTGTCTATTGTGGSCAGGATCATCTTGTCCGGAA 5413
Db ||||| 5342 CGCGTACTGCTGACGACAGGAGCGTGTCTATTGTGGSCAGGATCATCTTGTCCGGAG 5401
QY ||||| 5414 GCGAGTGTGTTCCGACAGGAGTCTCTACACAGGATTCGATGATGGAAGATG 5473
Db ||||| 5402 GCGAGTGTGTTATCCCGACAGGAGTCTCTTACCAGGATTCGATGATGGAAGATG 5461
QY ||||| 5474 TGCCTCACAACTTCTTACATCAGCAGGGAATCAGCTCGCGAGCAATCAACGCAAA 5533
Db ||||| 5462 TGTCTCACACCTCCCTTACATCAGCAAGGAATCAGCTCGCGAGCAATTCBAACAGAA 5521
QY ||||| 5534 GCGCTCGGGTGTGTCAAACGCCACCAAGCAAGCGGAGTGTGCTCCCGTGGTGA 5593
Db ||||| 5522 GCGCTCGGATTTGTCAAACAGCCACCAAGCAAGCGGAGTGTGCTCCCGTGGTGA 5581
QY ||||| 5594 GTCCAGTGGAGCCCTTGAGACCTTCTGGGGAGAGCATGTGGAATTCATCAGCGG 5653
Db ||||| 5582 GTCCAGTGGAGCCCTTGAGGCTCTTCTGGGCGAAACACATGTGGAATTCATCAGCGG 5641
QY ||||| 5654 AATACAGTACCTAGCAGGCTTATCACTCTGCTGGAACCCCGGATAGCATATTAT 5713
Db ||||| 5642 GATACAGTACTTGGCAGGCTTATCACTCTGCTTGGAAACCCCGGATAGCATATTAT 5701
QY ||||| 5714 GGCATTTACAGCTTCTTATCACTAGCCGCTCACCAACCAACACCCCTCTCTTTAAACAT 5773
Db ||||| 5702 GGCATTTACAGCTTCTTATCACTAGCCGCTCACCAACCAACACCCCTCTCTTTAAACAT 5761
QY ||||| 5774 CTTGGGGGATGGGTGGTGGCCAACTGCTCTCTCCAGCGCTGCTGAGCTTTCTGGG 5833
Db ||||| 5762 CTTGGGGGATGGGTGGTGGCCAACTGCTCTCTCCAGCGCTGCTGAGCTTTCTGGG 5821
QY ||||| 5834 CGCGGATCGCGGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGTCTCTGAGACAT 5893
Db ||||| 5822 CGCGGATCGCGGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGTCTCTGAGACAT 5881
QY ||||| 5894 CTTGGCGGCTATGGGCGAGGGTATGCGCGCACTCTGTGGCTTTTAAAGTTCATGAGCGG 5953
Db ||||| 5882 TCTGGCGGCTATGGGCGGGGTGGTGGCGCATCTGTTGGCTTTTAAAGTTCATGAGCGG 5941
QY ||||| 5954 CGAGTGGCTTCCACCGAGGACTGTGTCAACTTTATCTCCCTGCGCATCTCTCTCTGGTGC 6013
Db ||||| 5942 CGAGATGCTTCCACTGAGGATCTGTTAAATTTACTCTCCCTGCGCATCTCTCTCTGGCGC 6001
QY ||||| 6014 CTTGTGCTGGGTGCTGCGCAGCAATCTGCTCGGCACTGGGCGCGGAGAGG 6073
Db ||||| 6002 CTTGTGCTGGGTGCTGCGCAGCAATCTGCTCGGCACTGGGCGCGGAGAGG 6061
QY ||||| 6074 GCGTGTGAGTGGATGAACCGGCTGATAGGCTTCTGCTTCGCGGGGTAAACAGCTCTCCCC 6133

Db ||||| 6062 GCGTGTGAGTGGATGAACCGGCTGATAGAGCTTCGCTTCGCGGGGTAAACAGCTCTCCCC 6121
QY ||||| 6134 TACGCACTATGTGCTGAGAGCGAGCTGACAGCTGTCTACTCAGATCTCTCTTAGCT 6193
Db ||||| 6122 CACGCACTATGTGCTGAGAGCGAGCTGCGCGGGGTGTTACTCAGATCTCTCTCAGCT 6181
QY ||||| 6194 TACCATCACTCACTGTCTGAAGCGGCTCCACAGTGGATTAAATGAGGACTGCTCTTACGCC 6253
Db ||||| 6182 TACCATCACTCAGTTGCTGAAGAGGCTTCATCAGTGGATTAAATGAGGACTGCTCTCAGGCC 6241
QY ||||| 6254 ATGCTCCGCTCGTGGCTTAAGGATGTTTGGGATGATATGACCGTGTGAGTACTT 6313
Db ||||| 6242 TTGTTCCGCTCGTGGCTTAAAGGATGTTTGGGACTGGAATATGACCGTGTGAGTACTT 6301
QY ||||| 6314 CAAGACTGGCTCCAGTCCAACTCTGCGCGGTTACCGGGAGTCCCTTTCTCTCATG 6373
Db ||||| 6302 CAAGACTGGCTCCAGTCCAACTCTGCGCGGTTACCGGGACTCCCTTTCTCTCATG 6361
QY ||||| 6374 CCAACGCGGTTACAAGGAGTCTGCGCGGGGAGCGGCATCATGCAAAACACCTGCCCCATG 6433
Db ||||| 6362 CCAACGCGGTTACAAGGAGTCTGCGCGGGGATGGCATCATGCAAAACACCTGCCCCATG 6421
QY ||||| 6434 CGGAGCAGATCGCGGAGCATGTCAAAAACGTTTCCATGAGGATCTGTAGGCTTAGAAC 6493
Db ||||| 6422 TGGAGCAGATCACCGGACATGTCAAAAATGGCTCCATGAGGATTTGTGGGCCAAAAAC 6481
QY ||||| 6494 CTGAGCAACAGCTGGGACGGAAGTTCGCCCATCAACGCATACACACGGGACCTTGCAC 6553
Db ||||| 6482 CTGAGCAACAGCTGGGATGGAACATTTCCCATCAACGCATACACACGGGACCTTGCAC 6541
QY ||||| 6554 ACCCTCCCGCGCGCCAACTATTCCAGGCGCTATGCGGGGTGGTGTGAGGAGTACGT 6613
Db ||||| 6542 GCGCTCCCGCGCGGCACTATTCCAGGCGCTGTGGCGGGTGGTGTGAGGAGTACGT 6601
QY ||||| 6614 GGAGTTACGCTGTGGGGATTTCCACTAGTACGCGGCTAGCACCTGACCAACGTAAC 6673
Db ||||| 6602 GGAGTTACGCGGGTGGGGATTTCCACTAGCTGACGCGCATGACACTGACCAACGTAAC 6661
QY ||||| 6674 GTGCCATGCCAGGTTCCAGCGCCCGGATTTCCAGCGAGGTGGATGAGTACGGTTGCA 6733
Db ||||| 6662 ATGCCATGCCAGGTTCCAGCGCCCGGATTTCCAGCGAGGTGGATGAGTACGGTTGCA 6721
QY ||||| 6734 CAGTACCTCGCGGTCGAAACCTCTTCTACGGGAGGACGTCAGCTTCCAGGTCGGGCT 6793
Db ||||| 6722 CAGTATGCTCCAGTGTGCAAACTCTCTACGAGAGGAGTCTGTAATTCAGGTCGGGCT 6781
QY ||||| 6794 CAACCAATATCTGTGCGGTGCGAGTCCCATGCGAGCCCGAAACCGGACGTAACAGTGT 6853
Db ||||| 6782 CAACCAATATCTGTGCGGTGCGAGTCCCATGCGAGCCCGAAACCGGATGTGGCAGTGT 6841
QY ||||| 6854 TACTTCCATGCTCCAGTCCCTCCCATATTACAGCAGAGCGGCTAAGCTAGGCTGGC 6913
Db ||||| 6842 CACTTCCATGCTCCAGGCTCCCTCTCATATTACAGCAGAGCGGCTAAGCTAGGCTGGC 6901
QY ||||| 6914 TAGAGGCTCTCCCCCTCTTTAGCCAGTCTCATCAGTACGAGTGTGTCTGCGGCTCTTT 6973
Db ||||| 6902 CAGGGGCTCTCCCCCTCTTTGGCCAGTCTTTCAGTACGAGTGTGTCTGCGGCTCTTT 6961
QY ||||| 6974 GAAGCGCATGCACTACCCACCATGACTCCCGGAGCGCTGACCTCATCGAGGCGCAACT 7033
Db ||||| 6962 GAAGCGCATGCTACTACCCCATGACTCCCGGAGCGCTGACCTCATCGAGGCGCAACT 7021
QY ||||| 7034 CTTTGGCGGAGGATGGGCGGAAACATCACTCGCTGGGATCAGAGAAATAGGTAGT 7093
Db ||||| 7022 CTTTGGCGGAGGATGGGCGGAAACATCACTCGCTGGGATCAGAGAAATAGGTAGT 7081
QY ||||| 7094 AATTCTGAGCTCTTTTCGAAACCGCTTTCAGCGAGGGGATGAGAGGAGATATCGGTCGC 7153
Db ||||| 7082 AATCTGAGCTCTTTTCGATCCGATTCGGCGGTTGAGGATGAGAGGAAATATCGTCCC 7141
QY ||||| 7154 GCGGAGATCTTCGAAATCAAGAAATTCCTTCAGGCTTGCCTATGAGGCGAGCGCC 7213
Db ||||| 7142 GCGGAGATCTTCGAAATCAAGAAATTCCTTCAGGCTTGCCTATGAGGCGAGCGCC 7201

QY	7214	GGACTACAAATCCTCCACTGCTAGAGTCTGGAAAGGACCGGACTACGTCTCCCTCGGTGCT	7273
DB	7202	GGATTACACCCCTCCACTGCTAGAGTCTGGAAAGGACCGGACTACGTCTCCCTCGGTGCT	7261
QY	7274	ACACGAGTGCCTATGCGACCTACCAAGGCTCTCCAAATACACCTCCACGAGGAAGAG	7333
DB	7262	ACACGGGTGCGCTTTGCCATCTACCAAGGCGCCCAATACCACTCCACGAGGAAGAG	7321
QY	7334	GACGGTGTCTCAGCAGATCCCAATGTCTCTGCTTGGCGAGCTCGCCACTAAGAC	7393
DB	7322	GACGGTGTCTCAGCAGATCCCACTGCTCTCTGCTTGGCGAGCTCGCTACTAAGAC	7381
QY	7394	CTTTCGGTAGCTCGCGATCGTTCGGCGCTTGATAGCGGCACGCGACCGCCCTCTCGACCT	7453
DB	7382	CTTTCGGTAGCTCGCGATCGTTCGGCGCTTGATAGCGGCACGCGACCGCCCTCTCGACCT	7441
QY	7454	GGCTTCGACGACGGTGACAAAGATCCGAGTGTAGTCTCTCTGCTGCTGCTGCTGCTGCT	7513
DB	7442	GGCTTCGACGACGGTGACAAAGATCCGAGTGTAGTCTCTCTGCTGCTGCTGCTGCTGCT	7501
QY	7514	TGAAGGGAGCGCGGAGACCCCGATCTCAGGACGGGTCTTGGTCTACCGTGAAGGA	7573
DB	7502	CGAGGGAGACGAGGGAGACCCCGATCTCAGGACGGGTCTTGGTCTACCGTGAAGGA	7561
QY	7574	GGCTAGTGAGATGTCTGCTGCTCAATGTCTATATCTGCTGCTGCTGCTGCTGCTGCT	7633
DB	7562	AGCTGTGTGAGACGTCTGCTGCTCAATGTCTATATCTGCTGCTGCTGCTGCTGCTGCT	7621
QY	7634	GCCATGCGTTCGAGGAGAAAGTAAAGTGCCTCAACCCGTTGAGCAACTCTTTCGTGCG	7693
DB	7622	GCCATGCGTTCGAGGAGAGCAAGTTGCCATCAATCCGTTGAGCAACTCTTTCGTGCG	7681
QY	7694	TCACCACAAATGTCTACGCGCAACATCCGCGAGCGCAAGCTCCGCGAGAAAGGT	7753
DB	7682	TCACCACAGTATGTCTACTCCACACATCTCGCAGCGCAAGTCTGCGCGAGAAAGGT	7741
QY	7754	CACCTTTGACAGATTGCAAGTCTGGATGATCATATTCGCGGACGCTACTCAAGGAGATGA	7813
DB	7742	CACCTTTGACAGATTGCAAGTCTGGACGACCACTACCGGACGCTGCTCAAGGAGATGA	7801
QY	7814	GGGAGGCGTCCAGATTAGGCTTAAGCTTCTATCTATAGAGAGGCGCTCGAAGCTGAC	7873
DB	7802	GGGAGGCGTCCAGATTAGGCTTAAGCTTCTATCTATAGAGAGGCGCTCGAAGCTGAC	7861
QY	7874	GCCCCACATTCGGCCAAATCCAAATTTGGCTATGGGCAAGGACGCTCCGGAACCTATC	7933
DB	7862	GCCCCACATTCGGCCAAATCCAAATTTGGCTATGGGCAAGGACGCTCCGGAACCTATC	7921
QY	7934	CAGCAGGCGCTTAACCACTCCGCTCGCTGTGGGAGGACTTGTCTGGAAGACACTGAAC	7993
DB	7922	CAGCAGGCGCTTAACCACTCCGCTCGCTGTGGGAGGACTTGTCTGGAAGACACTGAAC	7981
QY	7994	ACCAATTGACACCACTCATGGCAAAAGTGAAGTTTCTGCGTCCAAACAGAGAAGG	8053
DB	7982	ACCAATTGATACCACTCATGGCAAAAGTGAAGTTTCTGCGTCCAAACAGAGAAGG	8041
QY	8054	AGGCGCAAGCGAGCTCGCTTATCGTATTCAGACCTGGAGTTCGTGTATGGAGAA	8113
DB	8042	AGGCGCAAGCGAGCTCGCTTATCGTATTCAGACCTGGGAGTTCGTGTATGGAGAA	8101
QY	8114	GATGGCGCTTTACGAGGTGCTCCACCTTCTCAGGCGGTGATGGGCTCTCATACGG	8173
DB	8102	GATGGCGCTTTACGAGGTGCTCCACCTTCTCAGGCGGTGATGGGCTCTCATACGG	8161
QY	8174	ATTTCAATACTCCCAACGACGCGGTGAGTTCCTGTTGAATACTTGGAAATCAAAGAA	8233
DB	8162	ATTTCAATACTCCCAACGACGCGGTGAGTTCCTGTTGAATACTTGGAAATCAAAGAA	8221
QY	8234	ATGCGCTATGGGCTTCTCATATGACACCGGCTGTTTGAATCAACGGTCACTGAGAGTGA	8293
DB	8222	ATGCGCTATGGGCTTCTCATATGACACCGGCTGTTTGAATCAACGGTCACTGAGATGA	8281

QY	8294	CATTCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGGAGCCAGACAGCG	8353
DB	8282	CATTCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGGAGCCAGACAGCG	8341
QY	8354	CATAAGGTGCGTTCACAGAGCGGCTTTTACATCGGGGGTCCCTGACTAACTCAAAAGGCA	8413
DB	8342	CATAAGGTGCGTTCACAGAGCGGCTTTTATGTTCGGGGGTCCCTGACTAACTCAAAAGGCA	8401
QY	8414	GAACTCGGTTTATCGCGGTGCGGCAAGTGGCGTCTGACGACTAGCTGCGGTAATAC	8473
DB	8402	GAACTCGGTTTATCGCGGTGCGGCAAGTGGCGTCTGACGACTAGCTGCGGTAATAC	8461
QY	8474	CCTCAGTGTACTTTGAAGGCCACTTGTCTGTTATCTGTAAAGCGCGGGAACCCAGGAGA	8533
DB	8462	CCTCAGTGTACTTTGAAGGCCACTTGTCTGTTATCTGTGAGAGTGGGGAACCCAGGAGA	8521
QY	8534	GATGCTCGTGAACGAGACGACCTTGTCTGTTATCTGTAAAGCGCGGGAACCCAGGAGA	8593
DB	8522	GATGCTCGTGAACGAGACGACCTTGTCTGTTATCTGTGAGAGTGGGGAACCCAGGAGA	8581
QY	8594	TGCGGCGGCTTACGAGCCCTTCAAGAGGCTATGACTAGGTATTCGCGCCCCCGGGGA	8653
DB	8582	TGCGGCGGCTTACGAGCCCTTCAAGAGGCTATGACTAGGTATTCGCGCCCCCGGGGA	8641
QY	8654	TCGCGCCCCAACAGAAATACGACTGGAGCTGATAACATCATGTTCTCTCAATGTGTCACT	8713
DB	8642	CCGCGCCCCAACAGAAATACGACTGGAGCTGATAACATCATGTTCTCTCAATGTGTCACT	8701
QY	8714	CGGCGAGATGCTCTGGCAAAAGGTTATCTACCTCACCGTGACCCGACACACCCCT	8773
DB	8702	CGGCGAGATGCTCTGGCAAAAGGTTATCTACCTCACCGTGACCCGACACACCCCT	8761
QY	8774	TGCACGCGCTGCTGGGAGACAGCTAGACACACTCCAATCAACTCTTGGCTAGGCAATAT	8833
DB	8762	CGCAGCGGCTGCTGGGAGACAGCTAGACACACTCCAATCAACTCTTGGCTAGGCAATAT	8821
QY	8834	CATCATGTTATGCGCCCACTTGGCAAGGATGATTTCTGATGACTCATCTTTCTCCAT	8893
DB	8822	CATCATGTTATGCGCCCACTTGGCAAGGATGATTTCTGATGACTCATCTTTCTCCAT	8881
QY	8894	CTTCTAGCTCAAGAGCAACTTGAAGGCTTGGATTTGTGAGTCACTCATCTTTCTCCAT	8953
DB	8882	CTTCTAGCTCAAGAGCAACTTGAAGGCTTGGATTTGTGAGTCACTCATCTTTCTCCAT	8941
QY	8954	CTCATGAGCCACTTGAACCTACTCAGATCATTTGAACGACTCCATGCTCTTAGCGCAT	9013
DB	8942	CTCATGAGCCACTTGAACCTACTCAGATCATTTGAACGACTCCATGCTCTTAGCGCAT	9001
QY	9014	TACACTCCACAGTACTCTCCAGTGAGATCAATAGGTTGGCTTTCATGCTCAGGAACT	9073
DB	9002	TTCACTCCACAGTACTCTCCAGTGAGATCAATAGGTTGGCTTTCATGCTCAGGAACT	9061
QY	9074	TGGGTTACCACTTGGCAACTTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTACT	9133
DB	9062	TGGGTTACCACTTGGCAACTTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGTACT	9121
QY	9134	GTCCCAAGGGGGAGGCGCCACTTGTGGCAGATACCTCTTTAACTTGGGAGTAAAGAC	9193
DB	9122	GTCCCAAGGGGGAGGCGCCACTTGTGGCAGATACCTCTTTAACTTGGGAGTAAAGAC	9181
QY	9194	CAAGCTTAACTCACTCAATCCCGCGGTCCAGCTGAGCTTGTCTGGCTGGTTCGT	9253
DB	9182	CAAGCTTAACTCACTCAATCCCGCGGTCCAGCTGAGCTTGTCTGGCTGGTTCGT	9241
QY	9254	CGCTGGTTACAGCGGGGAGACATATATCAGAGCTCTCTGTCGCGGACCCCTCTGTT	9313
DB	9242	TGCTGGTTACAAACGGGGGAGACATATATCAGAGCTCTCTGTCGCGGACCCCTCTGTT	9301
QY	9314	TCGCTGTGCT	9373
DB	9302	CATGTTGCT	9361
QY	9374	AACGGGAGCTAAACCACTCCAGGCTTAAAGCCATTTCTCTGTTTTTTTTTTTTT	9424

Db 9362 AACGGGAGCTAACCACTCCAGCCNATAGGCCATTCCTCTTTT 9412

RESULT 11
ADP88596
ID ADF88596 standard; DNA; 9413 BP.
XX
AC ADF88596;
XX
DT 26-FEB-2004 (first entry)
XX
DE Hepatitis C virus NS3 gene, SEQ ID No 5.
XX
KW liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity; ds; gene.
XX
OS Hepatitis C virus.
XX
PN JP2003210181-A.
XX
PD 29-JUL-2003.
XX
PF 30-MAY-2002; 2002JP-00158335.
XX
PR 16-NOV-2001; 2001JP-00352443.
XX
PA (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.
XX
DR WPI: 2003-819836/77.
DR P-PSDB; ADF88597.
XX
PT Diagnosing liver cancer, involves amplifying amino terminal region of
PT hepatitis-C virus gene using predetermined primer and determining
PT hepatitis-C virus in base sequence of obtained DNA fragment.
XX
PS Disclosure; SEQ ID NO 5; 36pp; Japanese.
XX
CC The invention relates to the novel testing method for diagnosing liver
CC cancer. The novel method comprises amplifying the amino terminal region
CC of a hepatitis-C virus NS3 gene using a predetermined primer and
CC determining the hepatitis-C virus in a base sequence of the obtained DNA
CC fragment. The novel testing method is useful for diagnosing liver cancer
CC and also used in a gene amplification technique, a clinical laboratory
CC test reagent, a polymerase chain reaction, a base sequence analysis and
CC genetic engineering. The method enables the detection of a hepatitis-C
CC virus having high carcinogenicity with high specificity. This
CC polynucleotide sequence represents the hepatitis-C virus NS3 gene of the
CC invention.
XX
SQ Sequence 9413 BP; 1886 A; 2816 C; 2694 G; 2017 T; 0 U; 0 Other;
Query Match 86.6%; Score 8311.8; DB 10; Length 9413;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 8724; Conservative 0; Mismatches 687; Indels 0; Gaps 0;
QY 14 TGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATCTACTGTCTTCACGCGAGAA 73
Db 2 TGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATCTACTGTCTTCACGCGAGAA 61
QY 74 AGCGTCTAGCCATGGCTTAGTATGAGTGTCTGAGCTCCAGACCCCTCCCGGG 133
Db 62 AGCGTCTAGCCATGGCTTAGTATGAGTGTCTGAGCTCCAGACCCCTCCCGGG 121
QY 134 AGAGCCATAGTGGTCTGCGGACCGGTGAGTACACCGGAATTCAGGACGACCGGGTCC 193
Db 122 AGAGCCATAGTGGTCTGCGGACCGGTGAGTACACCGGAATTCAGGACGACCGGGTCC 181
QY 194 TTCTTGGATCAACCGGCTCAATGCTGGAGATTTGGGGGTGCTCCCGCGAGACTGCTAG 253
Db 182 TTCTTGGATCAACCGGCTCAATGCTGGAGATTTGGGGGTGCTCCCGCGAGACTGCTAG 241
QY 254 CCGAGTAGTGTGGGTGCGGAAAGCGCTTGGGTACTGCTGATAGGTGCTTGGAGTG 313

Db 242 CCGAGTAGTGTGGTCCGAAAGCCCTTGTGGTACTCGCTGATAGGGTGTCTTGGAGTG 301
QY 314 CCCGGGAGGTCTCGTAGACCGTGCACCATAGCACGAATCTTAAACCTCAAAGAAAAC 373
Db 302 CCCGGGAGGTCTCGTAGACCGTGCATCATGAGCACAAATCTTAAACCTCAAAGAAAAC 361
QY 374 CAAAGCTAAACCAACCGCCGCCACAGGAGCTCAAGTTCCCGGGCGGTGTGATCGT 433
Db 362 CAAAGCTAAACCAACCGCCGCCACAGGAGCTTAAAGTTCCCGGGCGGTGTGATCGT 421
QY 434 TGGTGGAGTTTACCTGTGGCCGCGAGGGGCCACAGTTGGGTGGTCCGCGACTAGGAA 493
Db 422 TGGTGGAGTTTACCTGTGGCCGCGAGGGGCCACAGTTGGGTGGTGGCGGACTAGGAA 481
QY 494 GGCTTCCGAGCGGTCCAAACCTCGTGAAGGCGACAACTATCCAAAGGTCGCGCGACC 553
Db 482 GACTTCCGAGCGGTCCAAACCTCGTGAAGGCGACAACTATCCAAAGGTCGCGCGACC 541
QY 554 CGAGGCGAGGCTCGGGCTAGCCCGGGTACCTTTGGCCCTCTATGGCAATGAGGGCT 613
Db 542 CGAGGCTAGGACTCGGGCTAGCCCGGGTACCTTTGGCCCTCTATGGCAATGAGGGTAT 601
QY 614 GGGGTGGGCGAGGATGGCTCCCTGTCACCCCGGGTCTCGGGCTAGTTGGGGCCCCACAGA 673
Db 602 GGGGTGGGCGAGGATGGCTCCCTGTCACCCCGGGTCTCGGGCTAGTTGGGGCCCCACAGA 661
QY 674 CCCCGGCGTAGGTCCGCTAACTTTGGGTAAAGGTCAATGATACCTTTACATCGGCTTCGC 733
Db 662 CCCCGGCGTAGGTCCGCTAACTTTGGGTAAAGGTCAATGATACCTTTACATCGGCTTCGC 721
QY 734 CGATCTCATGGGTACATTCGCTCGTGGCGCCCTAGGGGGCGCTGCAGGGCTT 793
Db 722 CGACTCATGGGTACATTCGCTCGTGGCGCCCTAGGGGGCGCTGCAGGGCTT 781
QY 794 GGCACAGGTGTCGGGTTCGAGGAGCGGCTGAACTATGCAACAGGGAATTCGCCGG 853
Db 782 GGCACAGGTGTCGGGTTCGAGGAGCGGCTGAACTATGCAACAGGGAATTCGCCGG 841
QY 854 TTGCTCTTTCTATCTCTCTTTGGCTCTGCTGCTGCTGTTGACCATCCAGCTTCGC 913
Db 842 TTGCTCTTTCTATCTCTCTTTGGCTCTGCTGCTGCTGTTGACCATCCAGCTTCGC 901
QY 914 TTATGAAGTGGCAACGTGTCGGGATATACCATGTCAAGCAAGCTGCTCCAACTCAAG 973
Db 902 TTACGAGTGGCAACGTGTCGGGATATACCATGTCAAGCAAGCTGCTCCAACTCAAG 961
QY 974 CATGTGTATGAGGCGAGGAGCTATGATGCTACTCCCGGGTGGTGGCTGTGTCTCA 1033
Db 962 TATTGTGTATGAGGCGAGGAGCTATGATGCTACTCCCGGGTGGTGGCTGTGTCTCA 1021
QY 1034 GGAGGTAACAGCTCCGTTGCTGGGTAGCGCTCACTCCACGCTCGCGGGCCAGGAATGC 1093
Db 1022 GGAGGTAATTTCTCCGTTGCTGGGTAGCGCTCACTCCACGCTCGCGGGCCAGGAACAG 1081
QY 1094 CAGGCTCCCACTACGACAAATACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153
Db 1082 CAGCATCCCAACGACAAATACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
QY 1154 CTGCTCCGCTATGATGCTGGGGATCTCTGGGATCTATTTTCTGCTGCTGCTGCTGCT 1213
Db 1142 CTGTTCCGCTATGATGCTGGGGATCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCT 1201
QY 1214 CACTTCTCGCTCGCGGATGAGAGCTGCAAGCTGCAACTGCTCAATCTATCCCGG 1273
Db 1202 CACTTCTCGCTCGCGGATGAGAGCTGCAAGCTGCAACTGCTCAATCTATCCCGG 1261
QY 1274 CCATGTATCAGGTCAACGCTGCTGGGATATGATGATGAATGCTGCTCAACAGC 1333
Db 1262 CCAGTATCAGGTCAACGCTGCTGGGATATGATGATGAATGCTGCTCAACAGC 1321
QY 1334 CCTAGTGGTGTGCGAGTTGCTCCGATCCCAAGAGCTGCTGAGCATGCTGGGGGGG 1393
Db 1322 CCTAGTGGTGTGCGAGTTGCTCCGATCCCAAGAGCTGCTGAGCATGCTGGGGGGG 1381

Qy	1394	CCACTGGGGAGTCTTGCGCGGCTTGCCCTACTATTCCATGGTAGGAACTGGGCTTAAGGT	1453
Db	1382	CCACTGGGGTGTCTTAGCGGCGCTTGCCCTACTATTCCATGGTAGGAACTGGGCTTAAGGT	1441
Qy	1454	TCTGATTGTGGCTACTCTTTTGGCGCGCTTGACGGGAGACCCACACGACGGGAGGGT	1513
Db	1442	CTTTGATTGTGATCTACTCTTTTGGCGCTTGACGGGACACCCACGTCAGACGGGGAG	1501
Qy	1514	GGCGGGCCACACCACTCGGGTTACGTCCTCTTTTCTCATCTGGGGCGTCTCAGAAAT	1573
Db	1502	GGTAGCCTCAGACCCAGAGCCTCGTGTCTCGGCTCTCAAGGCCCATCTCAGAAAT	1561
Qy	1574	CCAGCTTGTGAATACCAACGGCAGCTGGCACATCAACAGGACTGGCCTTAAATTCGAATGA	1633
Db	1562	CCAACTCGTGAACACCAACGGCAGCTGGCACATCAACAGGACCGCTCTGAAATGCAATGA	1621
Qy	1634	CTCCCTCAAACCTGGGTTCTTTTGGCGGCTGTGTTTACGACACACAGTTCAACTCGTCCGG	1693
Db	1622	CTCCCTCAAACCTGGGTTCTTTTGGCGGCTGTGTTTACGACACACAGTTCAACGCGCTCGG	1681
Qy	1694	GTGCGGGAGCGCATGSCCAGCTGCGCCCATTTGACTGTTGTCGCCACGGGTGGGGCC	1753
Db	1682	GTGCGGAGCGCATGSCCAGCTGCGCCCATTTGACTGTTGTCGCCACGGGTGGGGTCC	1741
Qy	1754	CATCACCTACTAAAGCTTAAACAGCTCGGATCAGAGCCCTTATTGTCGCATTATACGCGC	1813
Db	1742	CATCACTCATGATATGCTTGAGAGCTCGGACACAGAGGCCATATTGTGCGACTACGCGC	1801
Qy	1814	TCGACCGTGTGGTTCGTACCCGGTCGACAGTGTGTGTGTCAGTGTATTGTTTCACCCC	1873
Db	1802	TCGACCGTGTGGTTCGTACCCGGTCGACAGTGTGTGTGTCAGTGTATTGTTTCACTCC	1861
Qy	1874	AAGCCCTTTGTGTGGGGACCAACGATCGTTCCGGTCCCTACGTATAGCTTGGGGGA	1933
Db	1862	GAGCCCTTTGTGTGGGGACCAACGATCGTTCCGGTCCCTACGTATAGCTTGGGGGA	1921
Qy	1934	GAATGAGACAGACTGATGCTCTCAACAAACGCGTCCGCGTCCCTACGTATAGCTTGGGGGA	1993
Db	1922	GAATGAGACAGACTGATGCTCTCAACAAACGCGTCCGCGTCCCTACGTATAGCTTGGGGGA	1981
Qy	1994	CTGTACATGGATGAATAGTACTGGGTTTCACTAAGAGCTGCGGAGTCCGCCGTGTAACT	2053
Db	1982	GTGCACTGGATGAACAGCACTGGGTTTCAACAGAGCTGCGGGGCCCTCCGTGCAACAT	2041
Qy	2054	CGGGGGGTGCGTAAACCGCACCTTTGATCTGCGCCCAACGGAATGCTCGGAGCACCCCA	2113
Db	2042	CGGGGGGTGCGCAACAAACCTTTGTCTGCGCCCAACGGAATGCTCGGAGCACCCCA	2101
Qy	2114	GGCTACTTACAAAAATGTGGCTCGGGGCCCTGGTTGACACCTTAGTGTGCTAGTACATA	2173
Db	2102	GGCCACTTACAAAAATGTGGCTCGGGGCCCTGGTTGACACCTTAGTGTGCTAGTACATA	2161
Qy	2174	CCCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTTCATCTTTAAGGTAGGAT	2233
Db	2162	CCCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTTCATCTTTAAGGTAGGAT	2221
Qy	2234	GTATGTGGGGCGTGGAGCACAGGCTCAATGCGCATGCAATTTGGAATCGAGAGAGCG	2293
Db	2222	GTATGTGGGGCGTGGAGCACAGGCTCAATGCTGCATGCAATTTGGAATCGAGAGAGCG	2281
Qy	2294	CTGTAACCTTGAGAGCACAGGATAGGTGAGAACTCAGCCCGTGTCTGTCTACACAGA	2353
Db	2282	CTGTGACTTGGAGCACAGGATAGGTGAGAACTCAGCCCGTGTCTGTCTACACAGA	2341
Qy	2354	GTGGCAGATCTGCCCTGTCTTTCACACCACTACCGGCTTTATCCACTGGTTGATCCA	2413
Db	2342	GTGGCAGATCTGCCCTGTCTTTCACACCACTACCGGCTTTATCCACTGGTTGATCCA	2401
Qy	2414	TCTCCATCAGAACACTCGTGGAGCTGCAATACCTGTAGCGGTGAGGGTCAGCGTTTGTCTC	2473
Db	2402	TCTTCACGGAAACACTCGTGGAGCTGCAATACCTGTAGCGGTGAGGGTCAGCGTTTGTCTC	2461

QY	2474	CTTTGCAATCAAATGGGAGTACATCTGTGCTCTTTTCTCTCTCTCTGCGACAGCGCGCGT	2533
DB	2462	CTTTGCAATCAAATGGGAGTATATCTCTGTGCTTTTCTCTCTCTGCGGACGCGCGGT	2521
QY	2534	GTGTGCTGCTTTGTGGATGATGCTGCTCATAGCCAGGCTGAGCCGCGCTTATGAGAACTT	2593
DB	2522	CTGTGCTGCTTTGTGGATGATGCTGCTATAGCCAGGCTGAGCCAGCTTATGAGAACTT	2581
QY	2594	GGTGTCTCAATGCGGCGTCCGTGGCCGGAGCGCATGTATATCTCTCTTTCTTTGTGTT	2653
DB	2582	GGTGTCTCAATGCGGCGTCTGTGGCCGGAGCGCATGGCTTCTCTCTCTCTCTGTT	2641
QY	2654	CTTCTGCGCGCTCGTATCAATTAAGGCGAGGCTGGCTCCTCGGGCGGCGTATGCTTTTA	2713
DB	2642	CTTCTGCGCGCTCGTATCAATTAAGGCGAGGCTGGCTCCTCGGGCGGCGTATGCTTTTA	2701
QY	2714	TGGCGTATGGCGCTGCTCTGCTCTACTGTGCGTTACCAACAAGCTTACGCTTGGGA	2773
DB	2702	TGGCGTATGGCGCTGCTCTGCTCTACTGTGCGTTACCAACAAGCTTACGCTTGGGA	2761
QY	2774	CGGGAGATGGCTGATCGTGTGGGGGCTCGGTTCTTGTAGGCTGTGTTCTTTGACCTT	2833
DB	2762	CGGAGATGGCTGATCGTGTGGGGGCTCGGTTCTTGTAGGCTGTGTTCTTTGACCTT	2821
QY	2834	GTCAACATACTCAAAGTGTCTTCTCACTAGGCTCATATGTGTGTACAAATCTTTATCAC	2893
DB	2822	GTCAACATACTCAAAGTGTCTTCTCACTAGGCTCATATGTGTGTACAAATCTTTATCAC	2881
QY	2894	CAGAGCGAGGGCGACATGCAAGTGTGGGTCCCCCCCCCTCAACGTTCCGGGAGGCGCGA	2953
DB	2882	CAGAGCGAGGGCGACATGCAAGTGTGGGTCCCCCCCCCTCAACGTTCCGGGAGGCGCGA	2941
QY	2954	TGCCATATCTCTCTCACTGCTGTGGGTTCAATCCAGAGTTAATTTTGTGACATCACAACCT	3013
DB	2942	TGCCATATCTCTCTTACATGCGCGCTCCATCCAGAGCTAATCTTTGTGACATCACAACCT	3001
QY	3014	CCTGCTCGCCATACTCGGCCCGCTCATGTGTCTCCAGGCTGCATAACGAGAGTGCCTGA	3073
DB	3002	CCTGCTCGCCATACTCGGTCGCTCATGTGTCTCCAGGCTGCATACTAGAGTGCCTGA	3061
QY	3074	CTTCTGCGCGCTCAAGGCTCATTCGTGCATGCTATGTTAGTGCAGAAAGTCCGCGGGG	3133
DB	3062	CTTTGTACGGCTCAGGGCTCATCGTGCATGCTATGTTAGTGCAGAAAGTCCGCTGAGG	3121
QY	3134	TCATTATGTCCAAATGGTCTTTCATGAAGCTGGCGCGCTGACAGGTACGTAGCTTTATAA	3193
DB	3122	CCACTATGTCCAAATGGCTTTCATGAAGCTGGCGCGCTGACAGGTACGTAGCTATGA	3181
QY	3194	CCATCTTACCCCACTGGGGACTGGGCCACCGGGGCTTACAGAGCTTTCGGTGGCGGT	3253
DB	3182	CCATCTTACCCCACTGGGGATTTGGGCCCAACCGGGGCTTACAGAGCTTTCGGTGGCGGT	3241
QY	3254	AGAGCCGTGCTCTTCTCCGCATGAGACCAAGGTCATCACTGGGGAGCAGACACCG	3313
DB	3242	AGAGCCGTGCTCTTCTCTGACATGAGACTAACTCATCACTGGGGGAGCAGACACCG	3301
QY	3314	TGCGTGTGGGACATCATCTTGGGCTTACCCGCTCTCCGCCGGAAGGGGGAAGAGATAT	3373
DB	3302	GGCGTGTGGGACATCATCTCTGGGCTCTACAGGCTCTCCGCCGGAAGGGGGAAGAGATCT	3361
QY	3374	TTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGGTGGCGACTCTTTCGCCCATCACGGC	3433
DB	3362	TCTAGGACCGGCGATAGTTTGTGAGAGCAGGGGTGGCGGCTCTCTTGCGCCATCACGGC	3421
QY	3434	CTACTCCCAACAAACGGGGGCGTACTTGGTTCATCATCACTAGGCTCACAGGCGGGGA	3493
DB	3422	CTATTCCCAACAAACGGGGGCGTCTTGGCTGTATCATCACTAGGCTCACAGGCTGGGA	3481
QY	3494	CAAGAACAGGTCGAAGGGGAGGTTCAAGTGGTTTCTACGCAACCAATCTTTCTCTGGC	3553
DB	3482	CAAGAACAGGTCGATGGGAGGTTCAAGTGGTCTCTCAACGCAACCAATCTTTCTCTGGC	3541
QY	3554	GACCTGCATCAACCGGCGTGTGCTGGACTGTCTACCATGGCGCTGGGTTCGAAACCCCTAGC	3613

||||| 7922 CAGCAGGGCCGTCACACCATCCGCTCCGTGCGGAGGACTTGTCTGGAAGACACTGAAAC 7981
||||| 7994 ACCAATTGACACCATCATCGCAAAAGTACAGTCTTTCGCTCCACACGAGAAAGG 8053
||||| 7982 ACCAATTGATACCATCATGGCAAAATAGAGTCTTCTCGCTCCACACGAGAAAGG 8041
||||| 8054 AGCGCGCAAGCAGCTCGCTTATCGTATTTCCAGACCTGGGAGTTCGTGTATGCGAGAA 8113
||||| 8042 AGCGCGCAAGCAGCTCGCTTATCGTATTTCCAGACCTGGGAGTACGTGTATGCGAGAA 8101
||||| 8114 GATGCGCTTTACGACGTGGTCTCCACCTTCTCAGCGCTGATGGGCTCCTCATACGG 8173
||||| 8102 GATGCGCTTTACGACGTGGTCTCCACCTTCTCAGCGCTGATGGGCTCCTCATACGG 8161
||||| 8174 ATTTCAATACTCCCGCAAGCAGGGTCCAGTTCCTGTGTAATACCTGGAAATCAAGNA 8233
||||| 8162 ATTTCAATACTCCCGCAAGCAGGGTCCAGTTCCTGTGTAATACCTGGAAATCAAGNA 8221
||||| 8234 ATGCGCTATGGGCTTCTCATATGACACCGCTGTTTGTGACTCAACGGTCACTGACAGTGA 8293
||||| 8222 ATGCGCTATGGGCTTCTCATATGACACCGCTGTTTGTGACTCAACGGTCACTGACAGTGA 8281
||||| 8294 CATTCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAGGCCAGACAGC 8353
||||| 8282 CATTCGTGTTGAGGAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAGGCCAGCAGC 8341
||||| 8354 CATAGGTCGCTCAGACGGGCTTTACATCGGGGTCCTGACTCACTCAAAAGGGA 8413
||||| 8342 CATAGGTCGCTCAGACGGGCTTTATGTCGGGGTCCCTGACTTAATTCGAAGGGGCA 8401
||||| 8414 GAACTCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGAGCTAGCTAGCTGGGTAAATAC 8473
||||| 8402 GAACTCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGAGCTAGCTAGCTGGGTAAATAC 8461
||||| 8474 CCTCACATGTTACTTGAAGGCCACTCGAGCTGTGAGCTGCAAAAGTCCAGGACTGCAC 8533
||||| 8462 CCTCACATGTTACTTGAAGGCCACTCGAGCTGTGAGCTGCAAAAGTCCAGGACTGCAC 8521
||||| 8534 GATGCTCGTGAACGAGACGACTTGTGTTATCTGTGAAGCGGGGAACCCAGGAGGA 8593
||||| 8522 GATGCTCGTGAACGAGACGACTTGTGTTATCTGTGAGAGTGGCGGAACCCAGGAGGA 8581
||||| 8594 TGGCGGGCCCTACAGAGCTTCAAGAGGCTTACGAGCTATGCTAGCTATTCGCGCCCCCGGGGA 8653
||||| 8582 TGGCGGGCCCTACAGAGCTTCAAGAGGCTTACGAGCTATGCTAGCTATTCGCGCCCCCGGGGA 8641
||||| 8654 TCGCGCCCCAACCAAGATACGACCTGGAGCTGATAACATCATGTTCTTCCAAATGTGTGCT 8713
||||| 8642 CCGCGCCCCAACCAAGATACGACCTGGAGCTGATAACATCATGTTCTTCCAAATGTGTGCT 8701
||||| 8714 CGCGCAGATGATCTGGCAAAAGGATATACCTCAGCGGTGACCCCAACCCCTT 8773
||||| 8702 CGCGCAGATGATCTGGCAAAAGGATATACCTCAGCGGTGACCCCAACCCCTT 8761
||||| 8774 TGCACGGCTCGCTGGGAGACAGTACACACACTCCAATCACTCTTGGCTAGGCAATAT 8833
||||| 8762 CGCACGGCTCGCTGGGAGACAGTATAGACACTCCAGTCAACTCTTGGCTAGGCAATAT 8821
||||| 8834 CATCATGTTATGCGCCCCACCTATGGGCAAGGATGTTCTGATGACTCACTTTTCTCCAT 8893
||||| 8822 CATCATGTTATGCGCCCCACCTATGGGCAAGGATGTTCTGATGACTCACTTTTCTCTAT 8881
||||| 8894 CTTTCTAGCTCAAGAGCACTTGAAAGCCCTGGATTTGTACATCTACGGGGCTTGCTA 8953
||||| 8882 CTTTCTAGCTCAGAGCACTTGAAAGCCCTGGATTTGTACATCTACGGGGCTTGTTA 8941
||||| 8954 CTCATTGAGCACTTGACCTACTCAGATCATTTGAACGACTCCATGGTCTTAGCGCAAT 9013
||||| 8942 CTCATTGAGCACTTGACCTACTCAGATCATTTGAACGACTCCATGGTCTTAGCGCAAT 9001
||||| 9014 TACATCCACAGTACTCTCCAGGTGAGATCAATAGGGTGGCTTATGCTCTCAGGAACT 9073
|||||

Db 9002 TTCACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTATGCTCTCAGAAACT 9061
Qy 9074 TGGGGTACCACCTTGGAACTCGAGACATCGGGCCAGAAAGTGTCCGCGCTAAAGCTACT 9133
Db 9062 TGGGGTACCACCTTGGAGTCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAAGCTACT 9121
Qy 9134 GTCCAGGGGGGAGGAGGCCACTTGTGGCAGATACCTCTTTAACTGGGGCAGTAAGGAC 9193
Db 9122 GTCCAGGGGGGAGGAGGCCACTTGTGGCAGATACCTCTTTAACTGGGGCAGTAAGGAC 9181
Qy 9194 CAAGCTTAAACTCACTCCAATCCCGGCGGCTCCAGCTGGACTGTGTGGCTGGTTCGT 9253
Db 9182 CAAGCTTAAACTCACTCCAATCCCGGCGGCTCCAGCTAGACTTGTCCGGCTGGTTCGT 9241
Qy 9254 CGCTGGTTTACAGCGGGGAGACATATATACAGAGCTGTCTGTCGCCAGCCCGCTGGT 9313
Db 9242 TGCTGGTTTACAGCGGGGAGACATATATACAGAGCTGTCTGTCGCCAGCCCGCTGGT 9301
Qy 9314 TCCGTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCATTTACCTGTCTCCCAACCGGATG 9373
Db 9302 CATGTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCATCTACCTGTCTCCCAACCGGATG 9361
Qy 9374 AACGGGAGCTAACCACTCCAGGCTTAAAGCAATTTCTGTGTGTGTGTGTGTGT 9424
Db 9362 AACGGGAGCTAACCACTCCAGGCTTAAAGCAATTTCTGTGTGTGTGTGTGTGT 9412

RESULT 12
AAL53723
ID AAL53723 standard; DNA; 9413 BP.
XX
AC AAL53723;
XX
DT 27-OCT-2003 (revised)
DT 07-FEB-2003 (first entry)
XX
Hepatitis C Virus DNA sequence SEQ ID No 17.
XX
Target RNA; target RNA.support-attached test compound; flow cytometry;
KW mass spectrometry; high-throughput screening; ds.
XX
Hepatitis C virus; Virus.
XX
WO200283837-A1.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011758.
XX
PR 11-APR-2001; 2001US-0282966P.
XX
PA (PTCT-) PTC THERAPEUTICS INC.
XX
PI Almaden NG;
XX
WPI; 2003-075534/07.
PT Identifying a test compound that binds to a target RNA molecule by
PT separating the detectably labeled target RNA.support-attached test
PT compound complex from uncomplexed target RNA molecules and test compounds
PT by flow cytometry.
XX
PS Example; Page 55-59; 131pp; English.
XX
The invention relates to a novel method for identifying a test compound
CC that binds to a target RNA molecule comprising separating the detectably
CC labeled target RNA.support-attached test compound complex from
CC uncomplexed target RNA molecules and test compounds. The separating
CC process is carried out by flow cytometry and determining a structure of
CC the type of test compound of the RNA.support-attached test compound
CC complex by mass spectrometry. The method is useful for high-throughput
CC screening of libraries of compounds to identify pharmaceutical leads.
CC This polynucleotide sequence represents a DNA sequence related to the

CC	detecting method of the invention. (Updated on 27-OCT-2003 to standardise OS field)	
XX		
SQ	Sequence 9413 BP; 1887 A; 2815 C; 2694 G; 2017 T; 0 U; 0 Other; Query Match 86.6%; Score 8310.2; DB 8; Length 9413; Best Local Similarity 92.7%; Pred. No. 0; Matches 8723; Conservative 0; Mismatches 688; Indels 0; Gaps 0;	
QY	14 TGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATCTACTGTCTTCACGCAGAA 73	974 CATTTGTATGAGGCGGACGTGATCATCTCCCGGTGCGTCCCTGTGTTCA 1033
DB	2 TGGGGCGACACTCCACCATGATATCACTCCCTGTGAGGAATCTACTGTCTTCACGCAGAA 61	962 TATTTGTATGAGGCGGACGATATGACACACCCCGGTGCGTCCCTGTGTCG 1021
QY	74 AGCGTCTAGCCATGGCGTTAGTATGATGTGTGTCAGCCTCCAGGACCCCTCCCGGG 133	1034 GGAGGTAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCGGCGGCAGAAATGC 1093
DB	62 AGCGTCTAGCCATGGCGTTAGTATGATGTGTGTCAGCCTCCAGGACCCCTCCCGGG 121	1022 GGAGAGTAATTTCTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCGGCGGCAGAAACAG 1081
QY	134 AGAGCCATAGTGTCTGCGAACCCTGAGTACACCGGAATTCACAGGACCGGCTCC 193	1094 CAGCGTCCCACTACGACAAATACGACGCGACGTGACATGCTGCTTGGGACCGCTGCTTT 1153
DB	122 AGAGCCATAGTGTCTGCGAACCCTGAGTACACCGGAATTCACAGGACCGGCTCC 181	1082 CAGCATCCCACTACGACAAATACGACGCGACGTGCAATTTGCTTGGGCGCGCTGCTCT 1141
QY	194 TTTCTTGGATCAACCCGCTCAATGCTGGAGATTGCGGCTGCCCGGAGACTGCTAG 253	1154 CTGCTCCGTATGTAGTGGGGATCTCTGCGGATCTATTTTCTCGTCTCCACGCTGTT 1213
DB	182 TTTCTTGGATCAACCCGCTCAATGCTGGAGATTGCGGCTGCCCGGAGACTGCTAG 241	1142 CTGTTCCGTATGTAGTGGGATCTCTGCGGATCCGTTTTCTCGTCTCCACGCTGTT 1201
QY	254 CCGAGTAGTGTGGGTGCGAAAGGCTTGTGTACTTGCCTGATAGGCTGTTGCGAGTG 313	1214 CACCTTCTCGCTCCCGCATGAGACAGTGCAGGATGCAACTGCTCTAATCTATCCCGG 1273
DB	242 CCGAGTAGTGTGGGTGCGAAAGGCTTGTGTACTTGCCTGATAGGCTGTTGCGAGTG 301	1202 CACCTTCTCACCTCGCGGTATGAGACGGTACAAGATTGCAATTTGCTAATCTATCCCGG 1261
QY	314 CCCCGGAGGTCTCTGAGACCGTGCACCATGAGACAGAAATCTTAAACCTCAAGAAAC 373	1274 CCATGTATCAGGTACCGCATGGCTTGGATATGATGAACTGATGAACTGCTACCTACACAGC 1333
DB	302 CCCCGGAGGTCTCTGAGACCGTGCATCATGAGACAAATCTTAAACCTCAAGAAAC 361	1262 CCACGTATCAGGTACCGCATGGCTTGGGATATGATGAACTGCTACCTACACAGGC 1321
QY	374 CAAACGTAAACCAACCCGCGCCACAGGACGTCAAGTTCCCGGCGGTGCTCAGATCGT 433	1334 CCTAGTGGTTCGACGTTGCTCCGGATCCCAAGCTGCTGGACATGCTGGCGGGGC 1393
DB	362 CAAACGTAAACCAACCCGCGCCACAGGACGTTAAGTTCCCGGCGGTGCTCAGATCGT 421	1322 CCTAGTGGTTCGACGTTGCTCCGGATCCCAAGCCGCTGCGGACATGCTGGCGGGGC 1381
QY	434 TGGTGAAGTTTACCTGTTGCGCGACGGGCCCCAGGTTGGTGTGCGCGACTAGGAA 493	1394 CCACTGGGAGTCTCGCGGGCTTGCCTACTATTTCCATGTTAGGAACTGGGCTAAGGT 1453
DB	422 TGGTGAAGTTTACCTGTTGCGCGACGGGCCCCAGGTTGGTGTGCGCGACTAGGAA 481	1382 CCACTGGGTTGCTTAGCGGGCTTGCCTACTATTTCCATGTTGGGAACTGGGCTAAGGT 1441
QY	494 GGCTTCCGAGCGGTGCGAACTCTGTGGAAGCGCAACCTATCCCAAAGGCTGCCGACC 553	1454 TCTGATTGTGGCGTACTCTTTTCCGGGCTTGAACGGGAGACCCACACGACGGGAGGCT 1513
DB	482 GACTTCCGAGCGGTGCGAACTCTGTGGAAGCGCAACCTATCCCAAAGGCTGCCGACC 541	1442 CTTGATTGTGATGCTACTCTTTGCTGGCTGTGACGGGCAACACCACTGTCAGGGGGAAG 1501
QY	554 CGAGGCGAGGCTTGGGCTCAGCCCGGGTACCTTTGGCCCTCTATGGCAATGAGGCGCT 613	1514 GGCGGCGCACACACCTCCGGGTTACGTCCTTTTCTCATCTGGGGCGTCTCAGAAAT 1573
DB	542 CGAGGTTAGGACCTGGGCTCAGCCCGGGTACCTTTGGCCCTCTATGGCAACGAGGAT 601	1502 GGTAGCTTCCAGCACCCAGAGCCTCTGCTGCTGGCTCTCAAGGGCCCATCTCAGAAAT 1561
QY	614 GGGTGGGACGATGGCTCTGTCAACCCCGGCTCCGGCTAGTTGGGCGCCACGGA 673	1574 CAGCTTGTGAATACCAACGGCAGCTGGCACATCAACAGGACCTGCCCTAAATTTGCAATGA 1633
DB	602 GGGTGGGACGATGGCTCTGTCAACCCCGGCTCCGGCTAGTTGGGCGCCACGGA 661	1562 CCAACTCGTGAACACCAACCGGACGCTGGACATCAACAGACCGCTCTGAATTTGCAATGA 1621
QY	674 CCCCAGGCTAGGTCGGTAACTTTGGGTAAGGTATCATATACCTTATCATCGGCTTCGC 733	1634 CTCCCTCCAAACTGGGTTCTTTGCGCGCTGTTTTAGCGCACAAAGTTTCAACTGTCGCGG 1693
DB	662 CCCCAGGCTAGGTCGGTAAATTTGGGTAAGGTATCATATACCTTATCATCGGCTTCGC 721	1622 CTCCCTCCAAACTGGGTTCTTTGCTGCGCTGTTCTACGCGCACAGAGTTTCAACGCGTCCGG 1681
QY	734 CGATCTCATGGGTACATTTCCGCTCGTCCGGCCCCCTTAGGGGCGCTGCCAGGGGCTT 793	1694 GTGCCGAGCGCATGCGCAGCTGCGCCGCCATTTGACTGTTGCGCCAGGCGTGGGCGCC 1753
DB	722 CGACTCATGGGTACATTTCCGCTGTGCGCGCCCCCTTAGGGGCGCTGCCAGGGGCTT 781	1682 GTGCCAGAGCGCATGCGTACGCTGCGCCGCCATCGATGATGCTCAGGGGTGGGGTCC 1741
QY	794 GGCACACGGTGTCCGGGTTTGGAGGACGGCGTGAACTATGCAACAGGGAATTCGCCCGG 853	1754 CATCACTATATAAGCTTAACGCTCAGGATCAGAGGCTTATTTGCTGGCATTAACGGCC 1813
DB	782 GGCACATGGTGTCCGGGTTCTGGAGGCGGCGTGAATATGCAACAGGGAATTCGCCCGG 841	1742 CATCACTCATGATATGCTTGAAGCTCAGGACCTCGGACCGGATATTTGCTGGCATTAACGGCC 1801
QY	854 TTGCTCTTTCTATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913	1814 TCAGACCGTGGTGTGCTGCTACCCGCTCGCAGGTGTGTGCTCCAGTGTATTTGTTTCACCCC 1873
DB	842 TTGCTCTTTCTATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901	1802 TCGACCGTGGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1861
QY	914 TTATGAAGTGGCAACGCTGTCGGGATATACCATGTACAGAACGACTGCTCCAACTCAAG 973	1874 AAGCCCTGTTGTGGTGGGACCAACCGATCGTTCGGTGTCCCTACGTATAGCTGGGGGA 1933
DB	902 TTACGAGGTGGCAACGCTGTCGGGATATACCATGTACAGAACGACTGCTCCAACTCAAG 961	1862 GAGCCCTGTTGTAGTGGGACGACCGATCGTTTTCGGCGCTCTACTAGTATAGCTGGGGGA 1921

Db 2042 CGGGGGGTGCGCAACAACACCTTGCTGCTGCCCCACGGATTGCTTCCGGGAAGCACCCCGA 2101
Qy 2114 GGCTACTTTACAAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGTGCTAGTAGACTA 2173
Db 2102 GGGCACTTACAAAAAGTGTGGCTCGGGGCCCTGGTTGACACCCAGGTGCATGGTTGACTA 2161
Qy 2174 CCATACAGGCTTTTGGCACTACCCCTGCACCTCTCAATTTTTCATCTTTAAAGTTAGGAT 2233
Db 2162 CCCATACAGGCTTCTGGCACTACCCCTGCACCTGCACTGTTAACTTTTACCGTCTTTTAAAGTTCAGGAT 2221
Qy 2234 GTATGTGGGGGGCGTGGAGCACAGGCTCAATGCCGATGCAATTTGSACTCGAGGAGAGCG 2293
Db 2222 GTATGTGGGGGGCGTGGAGCACAGGCTCAAATGCTGCAATTTGSACTCGAGGAGAGCG 2281
Qy 2294 CTGTAACTTGGAGGACAGGATAGTGCAGAACTCAGCCGCTGCTGCTGTCTACAAACAGA 2353
Db 2282 CTGTAACTTGGAGGACAGGATAGTGCAGAACTCAGCCGCTGCTGCTGTCTACAAACAGA 2341
Qy 2354 GTGGCAGATACTGCCCTGTGCTTTTCCACACCTTATCCGCTTTATCCACTGGTTTGTATCCA 2413
Db 2342 GTGGCAGATACTGCCCTGTGCTTTTCCACACCTTATCCGCTTTTATCCACTGGTTTGTATCCA 2401
Qy 2414 TCTCCATCAGAAACATCGTGGGACGTGCAATPACCTGTACGGTGTAGGGTCAAGCTTTTGTCTC 2473
Db 2402 TCTTCAACGGGAACATCGTGGACGTGCAATPACCTGTACGGTGTAGGGTGGCGAGTTGTCTC 2461
Qy 2474 CTTTGAATCAAAATGGAGTACATCTGTTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2533
Db 2462 CTTTGAATCAAAATGGAGTATATCTGTTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2521
Qy 2534 GTGTGCTGCTTGTGGATGATGCTGCTGATAGCCAGGCTGAGCGCGCTTATAGAACTT 2593
Db 2522 CTGTGCTGCTTGTGGATGATGCTGCTGATAGCCAGGCTGAGCGCACCTTATAGAACTT 2581
Qy 2594 GGTGTGCTTCAATGGCGGCTCGGTGCGCGGAGCGCATGGTATTTCTTCTTCTTCTTCTTCTTCT 2653
Db 2582 GGTGTGCTTCAATGGCGGCTCTGTGGCGGAGCGCATGGCTTCTTCTTCTTCTTCTTCTTCTTCT 2641
Qy 2654 CTTTCTGCGCGCTGCTTACATTAAGGGCAGGTGCGTCTTGGGGCGGCGTATGCTTTTATA 2713
Db 2642 CTTTCTGCGCGCTGCTTACATCAAAGGCAAGGCTGGTCTTCTTGGGGCGGCGATATGCTCTCTA 2701
Qy 2714 TGGCGTATGGCGCTGCT 2773
Db 2702 TGGCGTATGGCGCTGCT 2761
Qy 2774 CCGGAGATGGCTGCAATCGTGGGGGGTGGCTTCTTGTAGTCTGCTGATTTCTTGACCTT 2833
Db 2762 CCGAGAGATGGCTGCAATCGTGGGAGGCGGCTTTTGTAGTCTGCTGATTTCTTGACCTT 2821
Qy 2834 GTCAACATACTAAAGTGTCT 2893
Db 2822 GTCAACATACTAAAGTGTCT 2881
Qy 2894 CAGCGCGAGGCGCATCAAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2953
Db 2882 CAGCGCGAGGCGCATCAAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2941
Qy 2954 TGCATATCT 3013
Db 2942 TGCATATCT 3001
Qy 3014 CTTGTGCGCATCT 3073
Db 3002 CTTGTGCGCATCT 3061
Qy 3074 CTTTGTGCGGCTCAAGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3133
Db 3062 CTTTGTGCGGCTCAAGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3121
Qy 3134 TCATTATGTCCAAATGGTCTTATGAAGCTGGGCGGCTGACAGGTACGTACGTTTATATA 3193

Db 3122 CCATATGTCCAAATGGCTTTCATAAAGTGGCGGCTGACAGGTACGTATGATATGA 3181
Qy 3194 CCATCTTACCCCACTGCGGGAATCGGGCCCAAGCGGCTTACGAGACCTTGGCGGTGGCGGT 3253
Db 3182 CCATCTTACTCACTGCGGGATTTGGGCCCAAGCGGCTTACGAGACCTTGGCGGTGGCGGT 3241
Qy 3254 AGAGCCGCTGCTTCTTCTCGGCATGAGACCAAGGTCACTACCTTGGGGAGGAGACACGCG 3313
Db 3242 AGAGCCGCTGCTTCTTCTGACATGAGACTAAACTCATCACTTGGGGGAGACACGCG 3301
Qy 3314 TGGTGTGGGACATCATCTTGGGTCTACCCGTCTCCGCCGAAAGGGGAGAGATAT 3373
Db 3302 GCGGTGTGGGACATCATCTCGGTCTACCACTCTCCGCCGAAAGGGGAGAGATAT 3361
Qy 3374 TTTGGGACCGCTGATAGTCTCGAAGGCAAGGTGGGAGCTCTTGGCGCCCATCACGCG 3433
Db 3362 TCTAGGACCGGCCCATAGTTTGGAGAGCAGGGTGGCGGTCTCTTGGCGCTATCACGCG 3421
Qy 3434 CTACTCTCCCAACAAACCGGGGCTACTTTGGTTGATCATCACTAGCTTCAAGGCGGGGA 3493
Db 3422 CTATTTCCCAACAAACCGGGGCTCTTGGCTGTATCATCACTAGCTTCAAGGTCGGGA 3481
Qy 3494 CAAGAACAGGTTCGAAAGGGAGGTTCAAGTGGTTTCTACCGCAACAATCTTTCTTGGC 3553
Db 3482 CAAGAACAGGTTCGATGGGAGGTTCAAGTGGTTTCTTCCACCGCAAGCAATCTTTCTTGGC 3541
Qy 3554 GACCTGCATCAACCGGCTGCTGCACTGTACCATGGCGCTGCTCGAAGACCTTACG 3613
Db 3542 GACCTGCATCAATGGCGGTGTTGGACCGTCTACCATGGTGGCGCTCGAAGACCTTGGC 3601
Qy 3614 CGGTCCAAAGGTCCAATCAACCAATGTACCAATGTAGACCTTGGACCTCGTGGGTG 3673
Db 3602 CGGCCGAGAGGTCCAATCAACCAATGTACCAATGTAGACAGGCTTCTTGGGTG 3661
Qy 3674 GCAGCGCCCCCGGGGCGCTCCATGACACCAATGTCAGCTGTGCGAGCTCGGACCTTTA 3733
Db 3662 GCGGCGCCCCCGGGGCGCTCCATGACACCGTGCACCTGCGGAGCTCGGACCTTTA 3721
Qy 3734 CTTGCTCAGAGACATGCTGATGCTCATTCGCGTGGCGCGGCGAGCGAGCAGCGGGAG 3793
Db 3722 CTTGCTCAGAGGATGCTGATGCTGCTTCCGCTGGCGCGGCGGCGAGCAGCGGGGAG 3781
Qy 3794 TCTACTCTCCCCCAGGCGCTCTCTACTGAAAGGCTCTTCCGCGTGGTCTTCAATTTGCTTG 3853
Db 3782 CTTGCTTCCCCCAGGCGCTCTCTACTGAGGGCTCTTCCGCTGGAGCCTTCTTCTTG 3841
Qy 3854 CTTTGGGGGCGCTGCTGCGGCTCTTCCGCGTGTGTGCAACCGGGGGTGGCGAA 3913
Db 3842 CTTTGGGGGCGCTGCTGCTGCTTCTTCCGCGCTCTTCCGCGCTCTTCCGCGGCTTGGCGAA 3901
Qy 3914 GCGGTGACCTTACACCGCTTGTAGTCTATGGAACCTACCATGCGGTCTTCCGCTTCTCAC 3973
Db 3902 GCGGTGACCTTACACCGCTTGTAGTCTATGGAACCTACCATGCGGTCTTCCGCTTCTCAC 3961
Qy 3974 AGACAACTCAACCCCGGCTGTACCCGAGACATTTCAAAGTGGCACATCTTCCAGCTTCC 4033
Db 3962 AGACAACTCATCCCTCTCGGCGGTACCCGAAACATTTCAAAGTGGCACATTTACAGCTCC 4021
Qy 4034 TACTGGCAGCGGAGACCAAAAGTCCCGCTGCTATGCAAGCCCAAGGTTCAAGT 4093
Db 4022 CACTGGCAGCGGAGAGCACCAAAAGTCCCGCTGCTATGCAAGCCCAAGGTTCAAGT 4081
Qy 4094 GCTGCTCTGAACCGCTCGGTTCGCGCACCTTAGGTTTGGGGGCTATATCTCCAGGC 4153
Db 4082 GCTGCTCTTAAACCGCTCGGTTCGCGCACATTTGGGCTTGGAGGCTATATGTTCAAGGC 4141
Qy 4154 ACAGGATTCGACCTTAAACATCAGAACTGGGTTAAGGACCATTTACAGCGGGGCTTCCAT 4213
Db 4142 ACATGGCATCGAGCTTAACTCAGAACTGGGGTAGGACCATTCACACGGGCGGCCCAT 4201
Qy 4214 TACGTACTCCACTTATGGCAAGTTCCTTGGCGAGCGGTGGCTGTTCTTGGGGGCGCTATGA 4273
Db 4202 CACGTACTCCACTTATGCAAGTTCCTTGGCGAGCGGTGGATGCTCCGGGGGCGCTATGA 4261

QY	4274	CATCAATAATGTGAGTGGCACTCAACTGACTCGACTACCATCTTGGGCATCGGCAC	4333
Db	4262	CATCAATAATGTGAGTGAATGCCACTCACTGACTCGACTACCATCTTGGGCATCGGCAC	4321
QY	4334	AGTCTGGACCAAGCGGAGACGGCTGGAGCGGGCTGCTGCTGCTCGGCACCGCTACACC	4393
Db	4322	AGTCTGGATCAGGCAGAGACGGCTGGAGCGGGCTCGTCTGCTCGGCACCGGCACGCC	4381
QY	4394	TCCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAAATAGGCGCTGTCGAACAATGG	4453
Db	4382	TCCGGGATCGATCACCGTGCCACACCCCAATCGAGGAAGTGGCCCTGTCCAACACTGG	4441
QY	4454	AGAGATCCCTCTATGCGAAAGCCATCCCATTTGAGGCCATCAAGGGGGGAGGCGATCT	4513
Db	4442	AGAGATCCCTCTATGCGAAAGCCATCCCATTTGAGGCCATCAAGGGGGGAGGCGATCT	4501
QY	4514	CATTTTCTGCAATTCGAAGAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG	4573
Db	4502	CATTTTCTGCAATTCGAAGAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG	4561
QY	4574	ACTGAACGCTGTAGCATATTAACGGGGCTTTGATGTGCTCGTCAATACCGCTATCGGAG	4633
Db	4562	ACTGAACGCTGTAGCATATTAACGGGGCTTTGATGTGCTCGTCAATACCGCTATCGGAG	4621
QY	4634	CGTGGTTGCTGGCAACAGACGCTCTAATGACGGGTTTACCGGCGATTTTCACTCAGT	4693
Db	4622	CGTGGTTGCTGGCAACAGACGCTCTAATGACGGGTTTACCGGCGATTTTCACTCAGT	4681
QY	4694	GATCGACTGAATACATGTGTACCCAGACAGCTCGACTTCAGCTTTGGATCCCACTTCAC	4753
Db	4682	GATCGACTGAATACATGTGTACCCAGACAGCTCGACTTCAGCTTTGGATCCCACTTCAC	4741
QY	4754	CATTGAGACGACGACCGTCCCAAGACGCGGTGTGCGCTCGCAACGGCGAGGTAGAAC	4813
Db	4742	CATTGAGACGACGACCGTCCCAAGACGCGGTGTGCGCTCGCAACGGCGAGGTAGAAC	4801
QY	4814	TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAACGGCCCTCGGSCAT	4873
Db	4802	TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAACGGCCCTCGGSCAT	4861
QY	4874	GTTGCAATCTTTCGGTCTGCTGTGAGTGTATGACGCGGGTGTGCTTGGTATGAGCTCAC	4933
Db	4862	GTTGCAATCTTTCGGTCTGCTGTGAGTGTATGACGCGGGTGTGCTTGGTATGAGCTCAC	4921
QY	4934	GCCGCTGAGACCTCGGTTAGGTTGCGGGTTACTTAATAACACAGGGTTGCGCGCTG	4993
Db	4922	GCCGCTGAGACCTCGGTTAGGTTGCGGGTTACTTAATAACACAGGGTTGCGCGCTG	4981
QY	4994	CCAGGACCATCTGGAGTTCTGGGAGAGCGCTTTCACAGGCGCTCACCCACATAGATGCCCA	5053
Db	4982	CCAGGACCATCTGGAGTTCTGGGAGAGCGCTTTCACAGGCGCTCACCCACATAGATGCCCA	5041
QY	5054	CTTCCTGTCCAGACTAAACAGGAGGAGACAACTTCTTCACTGTGGGCATATCAAGC	5113
Db	5042	CTTCCTGTCCAGACTAAACAGGAGGAGACAACTTCTTCACTGTGGGCATATCAAGC	5101
QY	5114	TACAGTGTGCGCAGGGGTCAAGCTCACCTCCATCTGTTGGGACCAAAATGTGGAAGTGTCT	5173
Db	5102	TACAGTGTGCGCAGGGGTCAAGCTCACCTCCATCTGTTGGGACCAAAATGTGGAAGTGTCT	5161
QY	5174	CATACGGCTGAACCTACACTGACGCGGCAACCCCTGCTGTATAGGCTAGGAGCGGT	5233
Db	5162	CATACGGCTGAACCTACACTGACGCGGCAACCCCTGCTGTATAGGCTAGGAGCGGT	5221
QY	5234	CCAAATGAGGTCTATCTCACACCCCATAACTAAATACATCATGCGATGCATGTTCGGC	5293
Db	5222	CCAAATGAGGTCTATCTCACACCCCATTAACCAATACATCATGCGATGCATGTTCGGC	5281
QY	5294	TGACCTGAGGTTGCTCACTAGCACCTGGGTGCTGTAGCGGAGTCTTTCAGGCTTTGGC	5353
Db	5282	TGACCTGAGGTTGCTCACTAGCACCTGGGTGCTGTAGCGGAGTCTTTCAGGCTTTGGC	5341

QY	5354	CGCATACTGCTGACGACAGGAGTGTGGTCAATCTTGGGACGAGCATCATCTTCTCGGGAA	5413
Db	5342	CGCGTACTGCTGACGACAGGAGTGTGGTCAATCTTGGGACGAGCATCATCTTCTCGGGAG	5401
QY	5414	GCAGCTGTGCTTCCGACAGGAAAGTCTCTACAGAGAGTTCGATGAGATGGAAGAGTG	5473
Db	5402	GCAGCTGTATTCGACAGGAAAGTCTCTACAGAGAGTTCGATGAGATGGAAGAGTG	5461
QY	5474	TGCCTCACAACTTCTTATCATCGACGAGGAATGAGCTTCGCGAGCAATTTCAAGCAAA	5533
Db	5462	TGCTTTCACACCTTCTTATCATCGACGAGGAATGAGCTTCGCGAGCAATTTCAAGCAAA	5521
QY	5534	GGCGTTCGGGTTTTCGCAAAACCGCCACCAAGCAAGCGGAGGCTGCTCCCGTGTGGA	5593
Db	5522	GGCGTTCGGATTTGCTCAAAACAGCCACCAAGCAAGCGGAGGCTGCTCCCGTGTGGA	5581
QY	5594	GTCCAAGTGGCGGCTTGTAGACCTTCTGGGCAAGACATGTGGAATTTTCAATCAGCGG	5653
Db	5582	GTCCAAGTGGCGGCTTGTAGGCTTCTTGGGGCAAAACATGTGGAATTTTCAATCAGCGG	5641
QY	5654	AATACAGTACTAGCAGGCTTATCCACTCTGCTCGTGAACCCCGCGATAGCATCATTGAT	5713
Db	5642	GATACAGTACTTGGCAGGCTTATCCACTCTGCTCGTGAACCCCGCGATAGCATCATTGAT	5701
QY	5714	GGCAATTAACAGCTTCTATCACTAGCCGCTCACCAACCAAAACACCCCTCTCTTTAACAT	5773
Db	5702	GGCTTTTACAGCTCTATCACAGCCGCTCACCAACCAAAATACCCCTCTCTTTAACAT	5761
QY	5774	CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTCGCTCAGCTTTTCGTGGG	5833
Db	5762	CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTCGCTCAGCTTTTCGTGGG	5821
QY	5834	CGCCGCATCGCCGAGCGGCTGTGCGACATAGCTTGGGAAGTGTCTCGTGGACAT	5893
Db	5822	CGCCGCATCGCCGAGCGGCTGTGCGACATAGCTTGGGAAGTGTCTCGTGGACAT	5881
QY	5894	CTTGGCGGGCTATGGCGCAGGGGTAGCCGCGCACTCGTGGCCCTTTAAAGTTCATGACGG	5953
Db	5882	CTTGGCGGGCTATGGCGCAGGGGTAGCCGCGCACTCGTGGCCCTTTAAAGTTCATGACGG	5941
QY	5954	CGAGGTGCCCTCCACCGAGGACCTGTGCTTACTCTCTGCGCATCTCTCTCTGTGTGC	6013
Db	5942	CGAGGTGCCCTCCACCGAGGACCTGTGCTTACTTACTCTCTGCGCATCTCTCTCTGTGTGC	6001
QY	6014	CTTGTGCTCGGGTGTGTGCGGACCAATCTCTGCTGCGCACTGTGGGCGCGGAGAGGG	6073
Db	6002	CTTGTGCTCGGGTGTGTGCGGACCAATCTCTGCTGCGCACTGTGGGCGCGGAGAGGG	6061
QY	6074	GGCTGTGAGTGGATGNAACGGCTGTAGGCTTTCGCTTCGCGGGTAAACACGCTCTCCCC	6133
Db	6062	GGCTGTGAGTGGATGNAACGGCTGTAGGCTTTCGCTTCGCGGGTAAACACGCTCTCCCC	6121
QY	6134	TAGCAGTATGTGCTCAGAGCGACGCTGACGACGCTGTCACTCAGATCTCTCTAGCCT	6193
Db	6122	CAGCAGTATGTGCTCAGAGCGACGCTGACGCGCGCTGTACTCAGATCTCTCTAGCCT	6181
QY	6194	TACCATCACTCAACTGTGGAAGCGGCTCCACAGTGGATTAATGAGGAGTGTCTAGGCC	6253
Db	6182	TACCATCACTCACTGTGGAAGCGGCTTCATCAGTGGATTAATGAGGAGTGTCTAGGCC	6241
QY	6254	ATGCTCGGGTGTGGCTAAGGAGTGTGGGATGGATATGACCGGTGTGTGACTGACTT	6313
Db	6242	TTGTTCCGGCTGTGGCTAAGGAGTGTGGGATGGATATGACCGGTGTGTGACTGACTT	6301
QY	6314	CAAGACCTGCTCAGTCCCAACTCTCTCGCGGCTTACCGGAGTCTCTCTCTGTCTATG	6373
Db	6302	CAAGACCTGCTCAGTCCCAAGCTCTCTCGCGGCTTACCGGAGTCTCTCTCTGTCTATG	6361
QY	6374	CCAACGGGGTACAAAGGAGTCTGGCGGGGAGCGGATCATGCAAAACCACTGCCCCATG	6433
Db	6362	CCAACGGGGTACAAAGGAGTCTGGCGGGGAGTGGGATCATGCAAAACCACTGCCCCATG	6421
QY	6434	CGGAGACAGATCGCCGGACATGTCAAAACCGGTTTCCATGAGGATCTGTAGGGCTTAGAAC	6493

Db 6422 TGGAGCAGATACCGGACATGTCAAATGGCTCCATGAGATTGTTGGGCCAAAAC 6481
QY 6494 CTGCAGCAACAGTGGCGACGGAACGTTCCCATCAACGCATACACACGGGACCTTGGAC 6553
Db 6482 CTGCAGCAACAGTGGCGATGGAACATCCCCATCAACGCATACACACGGGCCCTGCAC 6541
QY 6554 ACCCTCCCGGGCGCCAACTATTCCAGGGCGCTATGGCGGGTGGTGTCTGAGGAGTACGT 6613
Db 6542 GCCCTCCCGAGCGCGAACTATTCCAGGGCGCTGTGGGGGTGGTGTCTGAGGAGTACGT 6601
QY 6614 GGAGGTACGCTGTGGGGGATTTCCACTACGTGACGGGCATGACCACTGACCAAGTAAA 6673
Db 6602 GGAGGTACGCGGGTGGGGGATTTCCACTACGTGACGGGCATGACCACTGACCAAGTAAA 6661
QY 6674 GTGCCCATGCCAGGTTCCGGCCCCCGAAATTTCTACGGAGGTGGATGGAGTGGGTTGCA 6733
Db 6662 ATGCCCATGCCAGGTTCCAGCCCCCTGAATTTTACGGAGGTGGATGGAGTGGGTTGCA 6721
QY 6734 CAGGTACGCTCGGGCGTCCAAACCTTTCTACGGGAGGACGTCACGTTCCAGGTCGGCT 6793
Db 6722 CAGGTATGCTCCAGTGTGCMAACCTCTCTACGAGGAGGTCGTATTTCCAGGTCGGCT 6781
QY 6794 CAACCAATACTGTGTGGGTCCAGCTCCCATGCGAGCGCCGACCGGAGCTAACAGTGT 6853
Db 6782 CAACCAATACTGTGTGGGTCCAGCTCCCATGCGAGCGCCGACCGGAGTGTGGCAGTGT 6841
QY 6854 TACTTTCCATGCTCACCGATCCCTCCCAATTAACAGCAGAGCGGTAAAGCTGAGCTGGC 6913
Db 6842 CACTTCCATGCTCACCGACCCCTCTCATATTACAGCAGAGCGGCAAGGCTAGGCTGGC 6901
QY 6914 TAGAGGCTCTCCCTCTTTAGCAGCTCATCAGCTAGCCAGTTGTCTGCGCCCTCTTT 6973
Db 6902 CAGGGGCTCTCCCTCTCTTGGCCAGCTCTTTCAGCTAGCCAGTTGTCTGCGCCCTCTTT 6961
QY 6974 GAAGCGGACATGCATACCCACCATGATCCCGGACCGCTGACCTCATCGAGGCCAACT 7033
Db 6962 GAAGCGGACATGTACTACCCCATCATGACTCCCGGACCGCTGACCTCATCGAGGCCAACT 7021
QY 7034 CTGTGTGGCGGAGGAGATGGGGGAAACATCACTCGCTGGAGTCAGAGATAAGGTAGT 7093
Db 7022 CTGTGTGGCGGAGGAGATGGGGCGGAAATACCCCTGTGGAGTCAGAAATAGAGTGT 7081
QY 7094 AATCTCGACTCTTTGCAACCGCTTTCAGCGGAGGGGATGAGAGGAGATATCGTGGC 7153
Db 7082 AATCTCGACTCTTTGATCCGATTCGGCGGCTGGAGATGAGAGGAAATATCGTCCC 7141
QY 7154 GCGGAGATCTTGGCAAAATCCAGGAAGTTCCCTCAGCGTTGCCCATATGGGCAAGGCC 7213
Db 7142 GCGGAGATCTTGGCAAAACCCAGGAAGTTTCCCGCAGCGTTGCCCATATGGGCAAGGCC 7201
QY 7214 GGACTACAATCTCCACTGTAGTCTCTGGAAGGACCGGACTACGTCCTCCGCTGGT 7273
Db 7202 GGATTACAACCTCCACTGTAGTCTCTGGAAGGACCGGACTACGTCCTCCCGCTGGT 7261
QY 7274 ACACGGATGCCCATTTGCCACCTACCAAGGCTCTCTCAATACCACCTCCACGGAGAAAG 7333
Db 7262 ACACGGGTGCCCTTTGGCATCTACCAAGGCCCCCCCAATACCACCTCCACGGAGAAAG 7321
QY 7334 GACGTTGTCTTGAAGAAATCCAAATGTGTCTTCTGCTTGGCGGAGTCCGCACTAAGAC 7393
Db 7322 GACGTTGTCTTGAAGAGTCCACCGTGTCTCTGCTTGGCGGAGTCCGCTACTAAGAC 7381
QY 7394 CTTTCGGTACCTCGGATCGTGGCGCTTGCATAGCGGACCGGCGCGCTTCTCGACT 7453
Db 7382 CTTTGGAGTCTCGGGTCTGGCGCTTGCATAGCGGCGACGGGACTGGCCCTTCCGATCA 7441
QY 7454 GGCCTCCGACGAGTGAACAAAGGATCCGACGTTGAGTGTGACTCTCTCATGCCGCCCT 7513
Db 7442 GGCCTCCGACGAGCGGACAAAGGATCCGACGTTGAGTGTGACTCTCTCATGCCGCCCT 7501
QY 7514 TGAAGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTACCGTGAAGTGA 7573

Db 7502 CGAGGAGAGCCAGGGGACCCCGACCTCAGCGACGGGTCTTGGTCTACCGTGAGCGGGGA 7561
QY 7574 GGCTAGTGAGGATGCTGCTGCTCAATGCTCTATACGTGGACAGCGGCCCTGATCAC 7633
Db 7562 AGCTGGTGAGGACGCTGCTGCTCAATGCTCTATACATGGACAGGTGCTTGTATCAC 7621
QY 7634 GCGATGCGCTCGGAGGAAAGTAAAGTGCCTCCATCAACCCGTTGAGCAACTCTTTGCTGG 7693
Db 7622 GCGATGCGCTCGGAGGAGCAAGTTGCCCATCAATCCGTTGAGCAACTCTTTGCTGG 7681
QY 7694 TCACCACAATGCTGTACGCCACAACATCCCGAGCGCAAGCCTCCGCGCAACAAAGT 7753
Db 7682 TCACCACAATGCTGTACTCCACAACATCTCGACGCCAAGTCTGCGGCAAGAAAGT 7741
QY 7754 CACCTTTGACAGATTGCAAGTCTCTGGATGATCAATTACCGGACGCTACTCAAGGAGATGAA 7813
Db 7742 CACCTTTGACAGATGCAAGTCTCTGGAGCACTATCCCGGACGCTGCTCAAGGAGATGAA 7801
QY 7814 GCGAAGGCGTCCACAGTTAAGCTTAAGCTTCTATATAGAGGAGGCTGCAAGCTGAC 7873
Db 7802 GCGAAGGCGTCCACAGTTAAGGCTTAGCTTCTATATAGAGGAGGCTGCAAACTGAA 7861
QY 7874 GCCCCCATTTCCGCCAAATCCAAATTTGGCTATGGGCAAGGACGTCGCGAACTATC 7933
Db 7862 GCCCCCATTTCCGCCAAATCCAAATTTGGCTATGGGCAAGGACGTCGCGAGCTATC 7921
QY 7934 CAGCAGGCGCTTAACACCATCCGCTCGCTGGGAGGACTTGTGGAGACACTGGAAC 7993
Db 7922 CAGCAGGCGCTCAACCATCCGCTCGCTGGGAGGACTTGTGGAGACACTGGAAC 7981
QY 7994 ACCAATTGACACCACTATGCGCAAAAGTGAAGTTTTCTGCTTCAACACAGAAAGG 8053
Db 7982 ACCAATTGATACCACTATGCGCAAAAGTGAAGTTTTCTGCGTCCAACACAGAAAGG 8041
QY 8054 AGCGCGCAAGCAGCTCGCTTATCGTATTCAGACCTGGAGTTCGTGTATGCGAGA 8113
Db 8042 AGCGCGCAAGCAGCTCGCTTATCGTATTCAGACCTGGGAGTACGTGTATGCGAGA 8101
QY 8114 GATGCGCTTTACGACGCTGCTCCACCTTCTCAGCGCTGATGGGCTCCTCATACGG 8173
Db 8102 GATGCGCTTTACGACGCTGCTCCACCTTCTCAGCGCTGATGGGCTCCTCATACGG 8161
QY 8174 ATTTCAATFACCTCCCAAGCAGCGGTCAGTTCTGTGTAATACCTGGAATCAAGAA 8233
Db 8162 ATTTCAAGTACTCTCTGGGCGAGCGGTGAGTTCCTGTGTAATACCTGGAATCAAGAA 8221
QY 8234 ATGCGCTATGGCTTCTCATATGACACCGCTGTTTGACTCAACGCTCACTGAGAGTGA 8293
Db 8222 ATGCGCTATGGGCTTCTCATATGACACCGCTGCTTTGACTCAACGCTCACTGAGATGA 8281
QY 8294 CATTGCTGTTGAGGAGTCAATTTTACCAATGTTGACTTGGCCCCCGAGGCCAGACAGC 8353
Db 8282 CATCGTACTGAGGAATCAATTTTACCAATGTTGACTTGGCCCCCGAGGCCAGCAGC 8341
QY 8354 CATAAGGTCGCTCAGACGCGCTTTTACATCGGGGCTCCCTGACTAACTCAAAAGGCA 8413
Db 8342 CATAAGGTCGCTCAGACGCGCTTTTATGTCGGGGGCTCCCTGACTAATTCGAAGGGCA 8401
QY 8414 GAACTGCGGTTATCGCGGTGCCGCGCAAGTGGGCTGTGACGACTAGTGGGTTAATAC 8473
Db 8402 GAACTGCGGTTATCGCGGTGCCGCGCAAGTGGGCTGTGACGACTAGTGGGCAACAC 8461
QY 8474 CCTCACATGTTACTTGAAGGCCACTGACGCTGTGAGCTGCAAGCTCCAGGACTGAC 8533
Db 8462 CCTCACATGTTACTTGAAGGCCACTGCGGCTGTGAGCTTGCATGCAAGCTCCAGGACTGAC 8521
QY 8534 GATGCTCGTGAACGAGACGACCTTTGCTGTTATCTGTGAAAGCGCGGAAACCCAGAGGA 8593
Db 8522 GATGCTCGTGAACGAGACGACCTTTGCTGTTATCTGTGAGAGTGGCGAAACCCAGAGGA 8581
QY 8594 TCGCGCGGCTTACGAGGCTTCAAGGAGTATGACTAGGTTATTCGCGCCCCCGGGGA 8653
Db 8582 TCGCGCGGCTTACGAGGCTTCAAGGAGTATGACTAGGTTATTCGCGCCCCCGGGGA 8641

QY 494 GGCTTCCGAGCGGTCCGAACCTCGTGAAGGCGCAACACTATCCCAAAGGCTCGCCGACC 553
DB 482 GACTTCCGAGCGGTCCGAACCTCGTGAAGGCGCAACCTATATCCCAAAGGCTCGCCGACC 541
QY 554 CGAGGCGAGGCGCTGGGCTCAGCCCGGTACCCCTTGGGCCCTCTATATGCAATGAGGGCTT 613
DB 542 CGAGGCTAGGACTCGGGCTCAGCCCGGTACCCCTTGGGCCCTCTATATGCAACAGAGGCTAT 601
QY 614 GGGGTGGCAGATGGCTCTGTGTACCCCGGGCTCCCGGCTAGTGTGGGCCCCACAGGA 673
DB 602 GGGGTGGCAGATGGCTCTGTGTACCCCGGTGGCTCTCGGCCCTAGTGTGGGCCCCACAGGA 661
QY 674 CCCCCGGGTAGTTCGGCTAACTTCGGTAAAGTCAATCGATACCTTACATGCGGCTTCGC 733
DB 662 CCCCCGGGTAGTTCGGCTAACTTCGGTAAAGTCAATCGATACCTTACATGCGGCTTCGC 721
QY 734 CGATCTCATGGGGTACATTCGGCTCGTCGGGCCCCCCCTAGGGGCGCTGCGAGGGCTT 793
DB 722 CGACCTCATGGGGTACATTCGGCTCGTCGGGCCCCCCCTAGGGGCGCTGCGAGGGCTT 781
QY 794 GGCACACGGGTGCCGGGTCTGGAGACGGCGTGAACCTATGCAACAGGGAATCTGCCCGG 853
DB 782 GGCACATGGGTGCCGGGTCTGGAGACGGCGTGAACCTATGCAACAGGGAATCTGCCCGG 841
QY 854 TTGCTCTTCTCTATCTTCTCTCTGCTCTGCTGTCTGCTTGTGACCATCCAGCTTCGCG 913
DB 842 TTGCTCTTCTCTATCTTCTCTTCTAGCTTTGCTGTCTGTTGTGACCATCCAGCTTCGCG 901
QY 914 TTATGAAGTGCACAACTGTCGGGATATACATGTACGAACTGCTGCTCAACTCAAG 973
DB 902 TTACGAGGTGCCAACTGTCGGGATATACATGTACGAACTGCTGCTCAACTCAAG 961
QY 974 CATTTGTATAGGCAGCGGACGTGATCATGCAATCTCCCGGGTGGCTGCTGTTGTTCA 1033
DB 962 TATTTGTATAGGCAGCGGACATGATCATGACACCCCGGGTGGCTGCTGCTGCTGCG 1021
QY 1034 GGAGGTTAACTGCTCCCGTGTGGGTAGCGCTCACTCCACGCTCGGGCCAGGAATGC 1093
DB 1022 GGAGAGTAATTTCTTCCGTTGCTGGGTAGCGCTCACTCCACGCTCGGGCCAGGAACAG 1081
QY 1094 CAGCGTCCCACTAGCAATAACGACGACGCTGCACTTTGCTGCTTGGGACGGCTGCTTT 1153
DB 1082 CAGCATCCCCACCAAGCAATAACGACGACGCTGCACTTTGCTGCTTGGGCGGCTGCTCT 1141
QY 1154 CTGCTCCGCTATGTAGTGGGGATCTCTGCGGATCTATTTTCTGCTCCCGCTCCAGCTGTT 1213
DB 1142 CTGTTCCGCTATGTAGTTGGGGATCTCTGCGGATCCGTTTTTCTGCTCTCCAGCTGTT 1201
QY 1214 CACCTTCTCGCTCGCGGCATGAGACAGTGCAGGACTGCAACTGCTCAATCTATCCCGG 1273
DB 1202 CACCTTCTCACCTCGCGGTATGAGACGATGCAAGATTGCAATCTATATCCCGG 1261
QY 1274 CCATGATCAGGTCAACCGATGGCTTGGGATATGATGAATGAACTGGTCACTTCAACAGC 1333
DB 1262 CCAGGTATCAGGTCAACCGATGGCTTGGGATATGATGAATGAACTGGTCACTTCAACAGC 1321
QY 1334 CTTAGTGTGTGCGAGTTGCTCGGATCCCAAGCTGTGCTGGACATGCTGGCGGGGGC 1393
DB 1322 CCTAGTGTGTGCGAGTACTCTCGGATCCCAAGCCGCTGCTGGACATGCTGGCGGGGGC 1381
QY 1394 CCACTGGGAGTCTTGGCGGGCTTGCCTACTATTTCATGTTAGGGAATCTGGGCTTAAGT 1453
DB 1382 CCACTGGGAGTCTTGGCGGGCTTGCCTACTATTTCATGTTAGGGAATCTGGGCTTAAGT 1441
QY 1454 TCTGATTTGGGGCTACTCTTTTGGCGGCTTGAACGGGAGACCCACAGCGGGGAGGCT 1513
DB 1442 CTTGATTTGATGCTACTCTTTGCTGGCGTTGACGGGCACACCCAGTGCAGGGGGAG 1501
QY 1514 GGCAGGCAACCACTCCGGGTACGCTCCCTTTTCTCATCTGGGGGCTCTCAGAAAT 1573
DB 1502 GGTAGCTCCAGCACCCAGAGCTCGTGTCTGGGCTCTCACAAGGCCCATCTCAGAAAT 1561

QY 1574 CCAGCTTTGTGAATACCAACCGCAGCTGGCAATCAACAGGACTGCCCTAAATTTGCAATGA 1633
DB 1562 CCAACTCGTGAACCAACGAGCGAGCTGGCAATCAACAGGACCGCTCTGAATTTGCAATGA 1621
QY 1634 CTCCTTCCAACTGGGTCTTTTGGCGGCTGTTTTTACGCACACAAGTTCAACTGCTCGCG 1693
DB 1622 CTCCTTCCAACTGGGTTCATTTGTCGCTGTTTTTACGCACACAAGTTTCAACGCGTCCGG 1681
QY 1694 GTGCCCGAGGCGCATGGCCAGCTCCCGCCCATTTGACTGTTTCGCCACAGGGTGGGGCC 1753
DB 1682 GTGCCAGAGCGCATGGTAGCTGCCGCCCATCGATGAGTTTCGCTCAGGGGTGGGGTCC 1741
QY 1754 CATCACTTATTAAGCCTAAACAGCTCGGATCAGAGGCTTTATTTGCTGGCAATTAACGCC 1813
DB 1742 CATCACTCATGATATGCTGAGAGCTCGGACAGAGGCCATATTTGCTGGCACTACGCGCC 1801
QY 1814 TCGACCGTGTGCTGCTGATCCCGGCTGCGAGGTGTGGTTCAGTGTATTTGTTTCAACCCC 1873
DB 1802 TCGACCGTGTGCGGATCGTGGCTGCGAGTGTGGTTCAGTGTATTTGCTTCACTCC 1861
QY 1874 AAGCCCTGTTGTTGGTGGGACCAACCGATCGTTCCGCTGCTCCTACGTATAGCTGGGGGA 1933
DB 1862 GAGCCCTGTTGTTAGTGGGACGACCGATCGTTTCGCGCTCCTACGTATAGCTGGGGGA 1921
QY 1934 GAATGAGACAGACGTGATGCTCTCTCAACAAACACGCGTCCGCCACAAAGGCAACTGTTTCGG 1993
DB 1922 GAATGAGACAGACGTGCTGCTACTTTAGCAACACGCGGCCGCTCAAGGCAACTGTTTGG 1981
QY 1994 CTGTACATGATGATAGTACTGCGGTTCATTAAGACGTGGGAGTTCGCCGTGTAAAT 2053
DB 1982 GTGCACGTGGATGAACAGCACTGGGTTCAACAAGACGTGCGGGGCCCTCCGTGCAACAT 2041
QY 2054 CGGGGGGTTCGTTAACCGACCTGATCTGCCGCCACGACTGCTTCGCGAAGCACCCCGA 2113
DB 2042 CGGGGGGTTCGCGAACAAACCTTGGTCTGCCACAGGATGCTTTCGGAAGCACCCCGA 2101
QY 2114 GGCTACTTACACAAATGTTGCTCGGGCCCTGGTTGACACCTAGTGCCTAGTAGACTA 2173
DB 2102 GGCCACTTACACAAAGTGTGGCTCGGGCCCTGGTTGACACCCAGGTGCATGGTTGACTA 2161
QY 2174 CCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCTTTTAAGGTTAGGAT 2233
DB 2162 CCCATACAGGCTTGGCACTACCCCTGCACTCTTAACTTTTAACTTTCAGGTCAGGAT 2221
QY 2234 GTATGTGGGGGTGGAGCACAGCTCAATGCGCATGCAATTTGACTCCAGGAGAGCG 2293
DB 2222 GTATGTGGGGGTGGAGCACAGCTCAATGCTGCAATGCAATTTGACTCGAGGAGAGCG 2281
QY 2294 CTGTAACCTTGGAGGACAGGATAGGTCAAGAACTCAGCCCGCTGCTGTCTTACACAGA 2353
DB 2282 CTGTGACTTGGAGGACAGGATAGGTCAAGAACTCAGCCCGCTGCTGTCTTACACAGA 2341
QY 2354 GTGGCAGATATCTGCCCTGTGTTTCAACCCCTACCGGCTTTATCCACTGTTTGTATCCA 2413
DB 2342 GTGGCAGATATCTGCCCTGTTCCTTCAACCCCTACCGGCTTCCACTGCTTGTATCCA 2401
QY 2414 TCTCCATCAGAACATCTGTGGACGTGCAATCTGCTACGCTGTAGGCTCAGCGTTGTCTC 2473
DB 2402 TCTTCCCGGAACATCTGTGGACGTGCAATCTGCTACCGGTATAGGCTCGGAGTTGTCTC 2461
QY 2474 CTTTCAATCAATGGGAGTACATCTGTTGCTTTTCTTCTTCCCTGGCAGACGCGCGCT 2533
DB 2462 CTTTCAATCAATGGGAGTATATCTGTTGCTTTTCTTCTTCCCTGGCAGACGCGCGCT 2521
QY 2534 GTGTGCTGCTGTGGATGATGCTGCTGATAGCCAGGCTGAGCGGCTTATAGAACTT 2593
DB 2522 GTGTGCTGCTGTGGATGATGCTGCTGATAGCCAGGCTGAGGCGACCTTATAGAACTT 2581
QY 2594 GGTGTGCTCAATCGGCGCTCGTGGCGGAGCATGCTATTTCTCTCTTCTTGTGTT 2653
DB 2582 GGTGTGCTCAATCGGCGCTCTGTGGCGGAGCATGCGCTTCTCTCTCTCTCTGTT 2641
QY 2654 CTTTCTGCGCCCTGTTACATTTAAGGGCAGGCTGCTCCTTGGGGCGGCTATGCTTTTTTA 2713

Db 2642 |||||CTTCTGCGCCGCTGGTACATCAAGCGCAGGCTGGTCCCTTGGGGCGGCATATGCTCTCTA 2701
Qy 2714 TGGCGTATGCGCGCGTCTCTGCTCTCTACTTGGGGTTACCACGAGCTTACGCGCTTGA 2773
Db 2702 TGGGTAATGCGCGCTGCTCTGCTCTTGTCTGGCTTACACGAGCTTATGCCATGA 2761
Qy 2774 CCGGGAGATGGCTGCATCGTGGGGGGTGGGTTCTTGTAGTCTGGTATCTTGGACCTT 2833
Db 2762 CCGAGAGATGGCTGCATCGTGGGAGCGCGGTTTTTGTAGGCTGGTACTCTTGGACCTT 2821
Qy 2834 GTCAACATCTACAAAGTGTCTCACTAGGCTCATATGCTGTGTACAACTTTATAC 2893
Db 2822 GTCAACATCTATAAGGTGTTCTCGCTAGGCTCATATGCTGTGTACAACTTTATAC 2881
Qy 2894 CAGAGCGAGCGGCACATGCAAGTGTGGGTCCCCCCCCCTCAACGTTTGGGGAGCGCGGA 2953
Db 2882 CAGAGCGAGCGGCACATTGCAAGTGTGGGTCCCCCCCCCTCAAGTTTGGGGAGCGCGGA 2941
Qy 2954 TGGCATCATCTCTCACTGCTGGGTTTCATCCAGAGTTAAATTTTGGACATCAACAACT 3013
Db 2942 TGGCATCATCTCTCACTGCGCGGTCCATCCAGAGCTAACTTTTGGACATCAACAACT 3001
Qy 3014 CTTGCTCGCGGCTCAAGGCTCATTTGCTGATGATGTTAGTGCAGAAAGTTCGCGCGGG 3133
Db 3002 CTTGCTCGCGGCTCAAGGCTCATTTGCTGATGATGTTAGTGCAGAAAGTTCGCGCGGG 3121
Qy 3134 TCATTATGTCAAATGCTTTCATGAAGTGTGGGCGCTGACAGGTACGTACGTTTATAA 3193
Db 3122 CCACTATGTCAAATGCTTTCATGAAGTGTGGGCGCTGACAGGTACGTACGTTATGA 3181
Qy 3194 CCATCTTACCCACTGCGGACTGGGCCACGCGGGCTACGAGACCTTGGGTTGGCGGT 3253
Db 3182 CCATCTTACTCCACTGCGGATTTGGGCCACGCGGGCTACGAGACCTTGGGTTGGCGGT 3241
Qy 3254 AGAGCCGCTGCTCTTCCGCGCATGAGACAAGGTCATCACCTTGGGGAGCAGACACCGC 3313
Db 3242 AGAGCCGCTGCTCTTCTGACATGGAGACTAACTCATCACCTTGGGGGAGCAGACCGC 3301
Qy 3314 TGGTGTGGGACATCATCTTGGGTCTACCGTCTCCGCGCGAAGGGGGAAGAGATATT 3373
Db 3302 GCGGTGTGGGACATCATCTCGGTCTACAGTCTCCGCGCGAAGGGGGAAGAGATACT 3361
Qy 3374 TTTGGGACCGCTGATAGTCTCGAAGGCAAGGTTGGGACTCCTTGGGCCCATCAAGC 3433
Db 3362 TCTAGGACCGCGCGATAGTTTGGAGAGCAGGGGTGGCGGCTCCTTGGCGCTATCAAGC 3421
Qy 3434 CTACTCCCAACAAAGCGGGCGTACTTGGTTGATCATCATAGCTTCAAGCGCGGGA 3493
Db 3422 CTATTCCCAACAAAGCGGGGCGTCTTGGCTGTATCATCACTAGCTTCAAGGTGGGA 3481
Qy 3494 CAAGAACAGGTGCAAGGGGAGGTTCAAGTGGTTTCTACCGCAACACAACTTTCTTGGC 3553
Db 3482 CAAGAACAGGTGCAATGGGAGGTTCAAGTGGTTTCTCCACCGCAACGCAACTTTCTTGGC 3541
Qy 3554 GACCTGCATCAAGCGCGTGTCTGGAATGTTACATGGCGCTGGCTCGAAGACCTTAGC 3613
Db 3542 GACCTGCATCAAGCGCGTGTGGACCGTCTACCAATGTACACCAATGTACACGAGACCTTGGC 3601
Qy 3614 CGGTCCAAAGGTCCAAATGACCCAAATGTACCAATGTACCAATGTACCAATGTACCAATGT 3673
Db 3602 CGGCCGAAGGTTCCAAATGACCCAAATGTACCAATGTACCAATGTACCAATGTACCAATGT 3661
Qy 3674 GCAGCGCCCCCGGGCGGCTCCATGACACCATGACAGCTGTGGAGCTCGGACCTTTA 3733
Db 3662 GCGCGCGCCCCCGGGCGGCTCCATGACACCGTGCACCTTGGGAGCTCGGACCTTTA 3721
Qy 3734 CTTGTCACGAGACATGCTGATGTCATTTCCGGTGGCGCGGAGCGCAGCAGCGGGAG 3793

Db 3722 CTTTGTACAGGCGATGCTGATGTCGTTCCGGTGGCGCGGGCGACAGCAGCGGGAG 3781
Qy 3794 TCTACTCTCCCGCAGCGCTCTCTACTCTGAAAGGCTCTCTCGGTGGTCCATGCTTTG 3853
Db 3782 CTTGCTTTCCCGCAGCGCTCTCTACTCTGAAAGGCTCTCTCGGTGGACCACTGCTTTG 3841
Qy 3854 CTTTCCGGGCGAGCTGCTGGGGTCTTCCGGGCTGCTGTGTGACCCGGGGGTCCGGA 3913
Db 3842 CTTTCCGGGCGAGCTTGTAGGCATCTTCCGGGCTGCTGTGTGACCCGGGGGTTCGGA 3901
Qy 3914 GCGGTGGACTTCAATACCGTTGAGTCTATGGAATCTACATCGGTCTCCGGTCTTTCAC 3973
Db 3902 GCGGTGGACTTCAATACCGTTGAGTCTATGGAATCTACATCGGTCTCCGGTCTTTCAC 3961
Qy 3974 AGCAACTCAACCCCCCGGCTGTACCGCAGACATTTCCAGTGGCACATCTGCACGCTCC 4033
Db 3962 AGCAACTCAATCCCCCTCCGCGGTACCCCAACATTTCCAGTGGCACATTTTACACGCTCC 4021
Qy 4034 TACTGCGAGCGGCAAGCAGACCAAGTGGCGGTGCGTATGCAAGCCCAAGGGTACAAGT 4093
Db 4022 CACTGCGAGCGGCAAGCAGACCAAGTGGCGGTGCTATGCAAGCCCAAGGGTACAAGT 4081
Qy 4094 GCTGCTCTGAAACCCCGTCCGTTGCGCCACCTTAAAGGTTTGGGGCGTATATGTCGAAGC 4153
Db 4082 GCTGCTCTTAAACCCCGTCCGTTGCGCCACATTTGGGCTTTGGAGCGTATATGTCGAAGC 4141
Qy 4154 ACAGGTATCGACCCCTAAACATCAGAACTGGGTGAAGCAACATTTACACGCGCGGCTCCAT 4213
Db 4142 ACATGATCTGAGCTTAAACATCAGAACTGGGGTGAAGCAACATTTACACGCGCGGCTCCAT 4201
Qy 4214 TAGTACTCCACCTATGGAAGTTCCTTGGCAGCGGTGGCTGTTTGGGGGCGGCTATGA 4273
Db 4202 CAGTACTCCACCTATTTGCAAGTTCCTTGGCAGCGGTGGATGCTCGGGGGCGGCTATGA 4261
Qy 4274 CATCAATAATATGATGAGTGCACCACTCACTGACTGACTTGGGCTATGCGGCATCGGCAC 4333
Db 4262 CATCAATAATATGATGAGTGCACCACTCACTGACTGACTTGGGCTATGCGGCATCGGCAC 4321
Qy 4334 AGTCTGGACCAAGCGGAGACGGCTGGAGCGGCTGCTGCTGCTCGCCACCGCTACAC 4393
Db 4322 AGTCTGGATCAGGACAGACGGCTGGAGCGGCTGCTGCTGCTCGCCACCGCAGCGC 4381
Qy 4394 TCCGGGATCGGTTACCGTGCACACCCCAATATCGAGGAATAGGCCCTGTCACCAATGG 4453
Db 4382 TCCGGGATCGATCACCGTGCACACCCCAATATCGAGGAATAGGCCCTGTCACCAATGG 4441
Qy 4454 AGAGATCCCTTCTATGGCAAGCGCATCCCATTTAGGCCATCAAGGGGGGGAGGCACTCT 4513
Db 4442 AGAGATCCCTTCTATGGCAAGCGCATCCCATTTAGGCCATCAAGGGGGGGAGGCACTCT 4501
Qy 4514 CATTTCTGCAATTCGAAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG 4573
Db 4502 CATTTCTGCAATTCGAAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG 4561
Qy 4574 ACTGAACGCTGTAGCATTTACCGGGCTTGTGATGCTCGCTCATACGCTATCGGGA 4633
Db 4562 ACTCAATGCTGTAGCGTATTTACCGGGCTCGATGCTCGCTCATACGCTATCGGGA 4621
Qy 4634 CGTGGTGTGCTGGCAACAGACGCTCTAATGAAGGGTTTCAACCGCGATTTTGAAGTCACT 4693
Db 4622 CGTGGTGTGCTGGCAACAGACGCTCTAATGAAGGGTTTACCGCGATTTTGAAGTCACT 4681
Qy 4694 GATGACTGCAATATGATGTCTACCCAGACAGTCTGAGTCTGAGTCTGAGTCTTTCAC 4753
Db 4682 GATGACTGCAACACATGTCACCCAGACAGTCTGATTTTCACTTTTGGATCCACCTTTCAC 4741
Qy 4754 CATTTGAGACGACGACCGTGGCGGAGAGCGGCTCGGCTCGCAACCGGCGAGGTAGAAC 4813
Db 4742 CATTTGAGACGACGACGCTGGCGGAGAGCGGCTCGGCTCGGAGGTAGAAC 4801
Qy 4814 TGGCAGGGGTAGGAGTGGCATCTTACAGTGTGTCACCTCCAGAGAAAGCGGCTCGGCT 4873
Db 4802 TGGCAGGGGTAGGAGTGGCATCTTACAGTGTGTCACCTCCAGAGAAAGCGGCTCGGCT 4861

Db	7022	CTGTGGCGCAGGAGATGGCGGGAACATCACCCGTGTGGAGTCAGAAAAATAAGTGGT	7081
Qy	7094	AATTCTGGGACTCTTTTCCAAACCGCTTACGCGGAGGGGGATGAGAGGGAGATATCCGTGCG	7153
Db	7082	AATCCTGGACTCTTTCGATCCGATTTCGGCGGTGGAGGATGAGAGGAAATATCCGTGCC	7141
Qy	7154	GGCGGAGATCTTCGGAATAATCCAGGAAGTTCCCTCAGCGTTCGCCATATGGCAGGCC	7213
Db	7142	GGCGGAGATCTTCGGAATAATCCAGGAAGTTCCGCCAGCGTTGCCCATATATGGCAGGCC	7201
Qy	7214	GGACTACAACTCCACTGCTAGAGTCCTGGAAGGACCGGACTACGTCCCTCCGTGGT	7273
Db	7202	GGATTACAAACCTTCCACTGCTAGAGTCTTGGAAAGAACCCGGAAGTACGTCCCTCCGTGGT	7261
Qy	7274	ACACGGATGCCATATGCCACTTACCAAGGCTCTCCAATACCACTCCACGGAAGAAGAG	7333
Db	7262	ACACGGGTGCCCTTTGGCATCTACCAAGGCCCCCCCAATACCACTCCACGGAAGAAGAG	7321
Qy	7334	GACGGTGTCTGACAGAAATCCAAATGTGTCTTGTGCTTTGGCGGAGTCGCCACTAAGAC	7393
Db	7322	GACGGTGTCTGACAGAGTCCACCGTGTCTTGTGCTTTGGCGGAGTCGCTACTAAGAC	7381
Qy	7394	CTTCGGTAGCTCCGGATCGTCGCGGTGTGATAGGGGACGGGACCGCCCTTCTGACCT	7453
Db	7382	CTTTGGCAGCTCCGGGTCTGTCGCGGTGTGACGCGGCACGGGACTGGCCCTCCCGATCA	7441
Qy	7454	GGCTTCGAGCAGCGTGACAAAGATCCGACGTTGAGTCCGTACTCTCTCCATGCCCCCCT	7513
Db	7442	GGCTTCGAGCAGCGGCACAAAGATCCGACGTTGAGTCCGTACTCTCTCCATGCCCCCCT	7501
Qy	7514	TGAAGGGAGCCGGGGACCCCGATCTCAGCGACGGGTCTTGTGCTACCGTGTAGTAGGA	7573
Db	7502	CGAGGGAGAGCCAGGGACCCCGACCTCAGCGACGGGTCTTGTGCTACCGTGTAGTAGGA	7561
Qy	7574	GGCTAGTAGAGTGTCTGCTGCTCAATGTCTTATACGTGGACAGCGCCCTGATCAC	7633
Db	7562	AGCTGTGTAGGACGTCTGCTGCTCAATGTCTTATATACATGGACAGGTGCCCTTGATCAC	7621
Qy	7634	GCCATGCGCTCGGAGGAAAGTAGCTGCCATCAACCCGTTGAGCAACTCTTGTGTCG	7693
Db	7622	GCCATGCGCTCGGAGGAGAGCAAGTTGCCCATCATCTCGTTGAGCAACTCTTGTGTCG	7681
Qy	7694	TCACCAACAATGCTTACGSCACAACATCCGCGAGCGCAAGCCTCCGCGAGAGAAGGT	7753
Db	7682	TCACCAAGTATGCTTACTTCCACAACATCTCGCAGCGCAAGTCTCGCGAGAGAAGGT	7741
Qy	7754	CACCTTTGACAGATGCAAGTCTTGATGATCATTCGCGGACGTACTCAAGGAGATGAA	7813
Db	7742	CACCTTTGACAGACTGCAAGTCTTGAGCGACCACTACCGGGACGTGCTCAAGGAGATGAA	7801
Qy	7814	GGCGAAGCGTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGGAGCGCTGCAAGCTGAC	7873
Db	7802	GGCGAAGCGTCCACAGTTAAGGCTTAGGCTTCTATCTATAGAGGAGCGCTGCAAACTGAC	7861
Qy	7874	GCCCCACATTCGCGCAAAATCCAAATTTGGCTATGGGCGAAAGGACGTCCGGAACCTATC	7933
Db	7862	GCCCCACATTCGCGCAAAATCCAAATTTGGCTACGGGCGAAGGACGTCCGGAACCTATC	7921
Qy	7934	CAGCAGGCGCTTAAACAATCCGCTCCGTGTGGAGGACCTTGCTGGAAGACACTGAAAC	7993
Db	7922	CAGCAGGCGCTTAAACAATCCGCTCCGTGTGGAGGACCTTGCTGGAAGACACTGAAAC	7981
Qy	7994	ACCAATTGACACCAACCATATGGCAAAAGTGAAGTTTCTGCGTCCAAACAGAGAGGG	8053
Db	7982	ACCAATTGATACCAACCATATGGCAAAATAAGGTTTCTGCGTCCAAACAGAGAGAGG	8041
Qy	8054	AGGCGCAAGCAGCTCCGCTTATCGTATCCAGACCTGGGAGTTCTGATGCGAGAA	8113
Db	8042	AGGCGCAAGCAGCTCCGCTTATCGTATTCAGACCTGGGGTACGTGTATGGAGAA	8101
Qy	8114	GATGGCCCTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGGGCTCTCATACGG	8173

Db	8102	GATGGCCCTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGGGCCCTCATACGG	8161
Qy	8174	ATTTCAATACTCCGCCAAGCAGCGGTGAGTTCCTGGTGAATACCTGGAATCAAGAA	8233
Db	8162	ATTCAGTACTCTCCTGGGCAGCGGTGAGTTCCTGGTGAATACCTGGAATCAAGAA	8221
Qy	8234	ATGCCCTATGGGCTTCTCATATGACACCGCTGTTTGTACTCAACGGTCACTGAGAGTGA	8293
Db	8222	ATGCCCTATGGGCTTCTCATATGACACCGCTGTTTGTACTCAACGGTCACTGAGATGA	8281
Qy	8294	CATTCTGTGTGAGGAGTCAATTTTACCAATGTTGTGACTTTGGCCCCCGAGGCCAGACGG	8353
Db	8282	CATTCTGTGTGAGGAGTCAATTTTACCAATGTTGTGACTTTGGCCCCCGAGGCCAGACGG	8341
Qy	8354	CATAAGGTCTGCTCAGAGAGCGCTTTACATCGGGGTCCCTGACTAACTCAAAAGGGCA	8413
Db	8342	CATAAGGTCTGCTCAGAGAGCGCTTTTATGTGGGGTCCCTGACTAACTCGAAGGGGCA	8401
Qy	8414	GAACTCGGTATTCGCGGTGCGCGCAAGTGGCGTGTGACGACTAGCTGCGGTAAATAC	8473
Db	8402	GAACTCGGTATTCGCGGTGCGCGCAAGTGGCGTGTGACGACTAGCTGCGGTAAATAC	8461
Qy	8474	CCTTCATGTTTACTTTGAAGGCCACTGACGCTGTGAGCTGCAAAAGCTCCAGACTGCAC	8533
Db	8462	CCTTCATGTTTACTTTGAAGGCCACTGCGGCTGTGAGCTGCAAAAGCTCCAGACTGCAC	8521
Qy	8534	GATGCTGTGAACGAGAGACGACTTGTCTGTTATCTGTGAAGCGCGGAAACCCAGAGGA	8593
Db	8522	GATGCTGTGAACGAGAGACGACTTGTCTGTTATCTGTGAGTGTGCGGAAACCCAGAGGA	8581
Qy	8594	TGCGGGCGCTTACGAGCCTTACGAGGCTGATGACTAGGTATTCGCGCCCCCGGGGA	8653
Db	8582	TGCGGGCGCTTACGAGCCTTACGAGGCTGATGACTAGGTATTCGCGCCCCCGGGGA	8641
Qy	8654	TCGCGCCCCAACCAAGATAAGACCTGAGGCTGATAACATCATGTTCTCCAATGTGTCAGT	8713
Db	8642	CCGCGCCCCAACCAAGATAAGACCTGAGGCTGATAACATCATGTTCTCCAATGTGTCAGT	8701
Qy	8714	CGCGCAGATGCAATCTTGGCAAAAGGATATCTACTCACCGGTGACCCGACCCGCCCT	8773
Db	8702	CGCGCAGATGCAATCTTGGCAAAAGGATATCTACTCACCGGTGACCCGACCCGCCCT	8761
Qy	8774	TGCAAGGGCTGCGTGGGAGACACTAGACACATCCCAATCAACTCTTGGCTAGGCAATAT	8833
Db	8762	CGCAAGGGCTGCGTGGGAGACACTAGACACATCCCAAGTCAACTCTTGGCTAGGCAATAT	8821
Qy	8834	CATCATGTATGCGCCCACTTATGGGCAAGGATTTCTGATGACTCACTTTTCTCCAT	8893
Db	8822	CATCATGTATGCGCCCACTTATGGGCAAGGATTTCTGATGACTCACTTTTCTCTAT	8881
Qy	8894	CCTTCTAGCTCAAGAGCACTTGAAAGCCCTGGATTTGTGAGTGTGAGTGTGAGTGTGCTA	8953
Db	8882	CCTTCTAGCTCAAGAGCACTTGAAAGCCCTGGATTTGTGAGTGTGAGTGTGAGTGTGCTA	8941
Qy	8954	CTCATGAGCCACTTGACCTTACCTCAGATCATTTGAAGGATTTCTGATGACTCACTTTTCTAGCGGAT	9013
Db	8942	CTCATGAGCCACTTGACCTTACCTCAGATCATTTGAAGGATTTCTGATGACTCACTTTTCTAGCGGAT	9001
Qy	9014	TACATCTCAGATTTACTTCTCAGGTGAGATCAATAGGTTGGCTTTCATGCTCTCAGGAACT	9073
Db	9002	TTCACTCCACAGTTACTTCTCAGGTGAGATCAATAGGTTGGCTTTCATGCTCTCAGGAACT	9061
Qy	9074	TGGGGTACCCCTTGGAACTTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTACT	9133
Db	9062	TGGGGTACCCCTTGGAGTCTGGAGACATCGGGCCAGAGTGTCCGCGCTAAGTACT	9121
Qy	9134	GTCCAGGGGGAGGGCCGCACTTGTGGCAGATACCTCTTTAACTTTGGGAGTAAAGAC	9193
Db	9122	GTCCAGGGGGAGGGCTGCACTTTCGCGCAAGTACTCTTTCACTTTGGGAGTAAAGAA	9181
Qy	9194	CNAGCTTAACTCACTCCAACTCCCGCGGCTCCAGCTGAGCTTGTCTGGCTGGTTCGT	9253
Db	9182	CNAGCTTAACTCACTCCAACTCCCGGCTGCGTCCAGCTAGACTTGTCTGGCTGGTTCGT	9241


```
QY 9254 CGCTGTTACAGCGGGGAGACATATATACAGCCCTGTCGTGCCCCGACCCCGCTGGTT 9313
D 9242 TGCTGTTACAAAGCGGGAGACATATATACAGCCCTGTCGTGCCCCGACCCCGCTGGTT 9301
QY 9314 TCCGTTGTGCTTACTCTCTTCTGTTAGGGTAGGCATTTTACTGCTCCCAACCGATG 9373
D 9302 CATGTTGTGCTTACTCTCTTCTGTTAGGGTAGGCATTTTACTGCTCCCAACCGATG 9361
QY 9374 AACGGGAGCTAACACATCCAGGCCCTTAAGCCATTTCCGTTTCTTTTCTTTT 9424
D 9362 AACGGGAGCTAACACATCCAGGCCCAATAGGCCATTTCCCTTTTCTTTTCTTTT 9412

RESULT 14
AAQ81559
ID AAQ81559 standard; DNA; 9413 BP.
XX
AC AAQ81559;
XX
DT 06-DEC-1995 (first entry)
XX
DE Hepatitis C virus helicase gene.
XX
KW Hepatitis C virus; HCV; non-A non-B; helicase gene; RNA helicase;
KW baculovirus; recombinant production; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 330..9362
FT repeat_unit /tag= a
FT repeat_unit /tag= c
FT repeat_unit /rpt_type= inverted
FT repeat_unit /tag= d
FT repeat_unit /rpt_type= inverted
XX
PN JP06319583-A.
XX
PD 22-NOV-1994.
XX
PF 18-SEP-1992; 92JP-00249241.
XX
PR 18-SEP-1992; 92JP-00249241.
XX
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX
PP WPI; 1995-040330/06.
XX
DR P-PSDB; AAR68864.
XX
of hepatitis C virus helicase gene in baculovirus - useful for large
PT scale prodn. of RNA helicase.
XX
PS Claim 1; Fig 1-4; 9pp; Japanese.
XX
CC AAQ81559 encodes AAR68864 hepatitis C virus (HCV) RNA helicase. The DNA
CC was used in the construction of an expression vector, which was used to
CC transform a baculovirus host. The transformed baculovirus could then be
CC used for the recombinant prodn. of HCV RNA helicase
XX
SQ Sequence 9413 BP; 1886 A; 2817 C; 2693 G; 2017 T; 0 U; 0 Other;

Query Match 86.6%; Score 8307; DB 2; Length 9413;
Best Local Similarity 92.7%; Pred. No. 0;
Matches 8721; Conservative 0; Mismatches 690; Indels 0; Gaps 0;

QY 14 TGGGGCGGACACTCCACCATGATCACTCCCTGTGAGGAATCTGCTTCCACGAGNA 73
D 2 TGGGGCGGACACTCCACCATGATCACTCCCTGTGAGGAATCTGCTTCCACGAGNA 61
QY 74 AGCGTCTAGCCATGGCGTTAGTATGATGTCGTGCGACCTCCAGGACCCCTCCCGG 133
```

```
Db
QY 62 AGCGTCTAGCCATGGCGTTAGTATGATGTCGTGCGACCTCCAGGACCCCTCCCGG 121
134 AGAGCCATAGTGGTCTCGGGAACCGGTAGTACACCGGAATTGCCAGGACACCGGTTCC 193
122 AGAGCCATAGTGGTCTCGGGAACCGGTAGTACACCGGAATTGCCAGGACACCGGTTCC 181
194 TTTCTTGGATCAACCCCTCAATGCCCTGGAGATTTGGGCGGTGCCCGCGAGACTGCTAG 253
182 TTTCTTGGATCAACCGCTCAATGCCCTGGAGATTTGGGCGGTGCCCGCGAGACTGCTAG 241
254 CCGAGTAGTGTGGGTCCGGAAGCCCTTGTGTACTCTGCTGATAGGGTGTCTGGAGTG 313
242 CCGAGTAGTGTGGGTCCGGAAGCCCTTGTGTACTCTGCTGATAGGGTGTCTGGAGTG 301
314 CCCGGGAGGTCTCGTAGACCGTGCACATGAGCAGCAATCCTAAACCTCAAGAAAC 373
302 CCCGGGAGGTCTCGTAGACCGTGCATGAGCAGCAATCCTAAACCTCAAGAAAC 361
374 CAAACGTAAACCAACCGCGCCGCCACAGGAGCTCAAGTTCCCGGGCGGTGTCAGATCGT 433
362 CAAACGTAAACCAACCGCGCCGCCACAGGAGCTTAAAGTTCCCGGGCGGTGTCAGATCGT 421
434 TGTGTGAGTTTACCTGTTGCCGCGCAGGGGCCCCAGAGTTGGGTGTGCGCGACTAGGAA 493
422 TGTGTGAGTTTACCTGTTGCCGCGCAGGGGCCCCAGAGTTGGGTGTGCGCGACTAGGAA 481
494 GGCTTCCGAGCGGTCCGCAACCTCGTGAAGGCGACCACTATCCAAAGGCTCCGCCGACC 553
482 GACTTCCGAGCGGTCCGCAACCTCGTGAAGGCGACCACTATCCCAAGGCTCCGCCGACC 541
554 CGAGGCGAGGCGCTGGGCTCAGCCCGGCTAGCTTGGGCCCTCTATGGCAATGAGGGCT 613
542 CGAGGCTAGGACCTGGGCTCAGCCCGGCTAGCTTGGGCCCTCTATGGCAATGAGGGCT 601
614 GGGTGGGCGAGGATGGCTCTGTCAACCCCGGCGCTCCCGGCTAGTTGGGGCGCCACGGA 673
602 GGGTGGGCGAGGATGGCTCTGTCAACCCCGGCGCTCTCGGCTAGTTGGGGCGCCACAGA 661
674 CCCCAGGCTAGTTCGGTAACTTGGGTAGGTTCATGATACCTTACATCGGCTTCGC 733
662 CCCCAGGCTAGTTCGGTAACTTGGGTAGGTTCATGATACCTTACATCGGCTTCGC 721
734 CGATCTCATGGGTACATTCGCTCGTGGGCGCCCTCCTAGGGGCGCTGTCAGGGCTT 793
722 CGACTCATGGGTACATTCGCTTGTGCGGCGCCCTCCTAGGGGCGCTGTCAGGGCTT 781
794 GGCACAGGCTTCGGGTCTGGAGGACGGCGTGAATATGCAACAGGGAACCTTGCCTGG 853
782 GGCACAGGCTTCGGGTCTGGAGGACGGCGTGAATATGCAACAGGGAACCTTGCCTGG 841
854 TTGCTCTTCTCATCTTCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
842 TTGCTCTTCTCATCTTCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
914 TTATGAGTGGCAACCGTTCGGGATATACATGTACGAACTGCTCAACTCAAG 973
902 TTACGAGTGGCAACCGTTCGGGATATACATGTACGAACTGCTCAACTCAAG 961
974 CATGTGTATGAGGACGGACGTGATCATGATCTCCCGGCTGCTGCTGCTGCTGCTGCTGCT 1033
962 TATTGTGTATGAGGACGGGACATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCT 1021
1034 GGAGGTTAACAGTCCCGTTCGGGTAGCGCTCACTCCAGCTCCGCGGCGAGGAATGC 1093
1022 GGAGGTTAACTTCTCCCGTTCGGGTAGCGCTCACTCCAGCTCCGCGGCGAGGAACAG 1081
1094 CAGCTGCCCACTACGACCAATACGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153
1082 CAGCATCCCAACGACCAATACGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
1154 CTGCTCCGCTATGATCGTGGGGATCTCTGCGGATCTATTTTCTGCTGCTGCTGCTGCTGCT 1213
```


Db 5522 GCGGCTCGGATGCTGCAAAACAGCCACCAAGCAAGCGGAGGCTGTGCTCCCGTGGTGA 5581
Qy 5594 GTCCAAGTGGCAGAGCCCTTTAGACCTTTGAGACCTTTGGCGAAGCAATGTGGAATTTTCATCAGCGG 5653
Db 5582 GTCCAAGTGGCAGAGCCCTTTAGGCTCTTCTGGCGGAAACACATGTGGAATTTTCATCAGCGG 5641
Qy 5654 AATACAGTACCTAGCAGGCTTATCCACTCTGCTTGGAAACCCGCGATAGCATATTGAT 5713
Db 5642 GATACAGTACTTTGGCAGGCTTATCCACTCTGCTTGGAAACCCGCGATAGCATATTGAT 5701
Qy 5714 GGCATTTACAGCTTCTATCAGTAGCCGCTTACACCCCAAAACACACCTCTCTGTTTAAACAT 5773
Db 5702 GGCATTTACAGCTTCTATCAGCCGCTTACACCCCAAAATACCTCTCTGTTTAAACAT 5761
Qy 5774 CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGCGTCAAGCTTTGCTGGG 5833
Db 5762 CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCCGCCAGCGCTTCCGCTTTCTGGG 5821
Qy 5834 CGCGGCTATCGCGAGCGGCTGTGGAGCATAGGCTTGGGAAGGTGCTCGTGACAT 5893
Db 5822 CGCGGCTATGGCGGTGCGGCGGTGGAGCATAGGTCTCGGGAAGGTACTTGTGGACAT 5881
Qy 5894 CTTGGCGGCTATGGGCGAGGGGTAGCCGCGCACTCGTGGCTTTAAAGTCAATAGCGG 5953
Db 5882 TCTGGCGGCTATGGGCGGCGGTGGCTGGCGCACTCGTGGCTTTAAAGTCAATAGCGG 5941
Qy 5954 CGAGGTGCCCTCCACCGAGGACCTGGTCAACTTACTCCCTGCCATCTCTCTCTGCTGC 6013
Db 5942 CGAGATGCCCTCCACTGAGGATCTGGTAAATTTACTCCCTGCCATCTCTCTCTGGCGC 6001
Qy 6014 CTTGGTGTGCGGGTCTGTGCGCAGCAATATCTGCTGCGCATAGTGGGCCCGGAGAGG 6073
Db 6002 CTTGGTGTGCGGGTCTGTGCGCAGCAATATCTGCTGCGCATAGTGGGCCCGGAGAGG 6061
Qy 6074 GGCCTGCAAGTGAATGAACCGGCTGATAGGCTTTCGCTTCCGCGGTAAACCAAGTCTCCCC 6133
Db 6062 GGCCTGCAAGTGAATGAACCGGCTGATAGGCTTTCGCTTCCGCGGTAAACCAAGTCTCCCC 6121
Qy 6134 TAGCCTATGTGCTGAGAGCAGCTGACGACGCTGTCACTCAGATCTCTCTAGCCT 6193
Db 6122 CAGCCTATGTGCGCGAGAGCGAGCGCGCGGCGGTGTACTCAGATCTCTCCAGCCT 6181
Qy 6194 TACCATCACTCAATCTGTAAGCGGCTCCACAGTGGATTAATGAGAGTGTCTTAGGCC 6253
Db 6182 TACCATCACTCAGTTGCTGAAGAGGCTTCACTCAGTGGATTAATGAGAGTGTCTCCAGGCC 6241
Qy 6254 ATGCTCCGCTCGTGGCTAAGGATGTTTGGGATTTGATATGACCGGTGTGACTGACTT 6313
Db 6242 TTGTTCCGGCTCGTGGCTAAGGATGTTTGGGACTGGATATGACCGGTGTGAGTGACTT 6301
Qy 6314 CAAGCTGGCTCCAGTCCAAACTCTCTGCGCGGTACCGGAGTCCCTTTCTCTGTCATG 6373
Db 6302 CAAGCTTTGGCTCAGTTCCAAGCTCTCTGCGCGGTACCGGACTCTCTTTCTGTCATG 6361
Qy 6374 CCAACGCGGTACAAAGGAGTCTGGCGGGGAGCGGATCATGCAAAACCACTGCCCCATG 6433
Db 6362 CCAACGCGGTACAAAGGAGTCTGGCGGGGATGGCATCATGCAAAACCACTGCCCCATG 6421
Qy 6434 CGGAGCAGATGCCCGGACATGTCAAAAACGGTTTCATAGAGTCTGTAGGGCTTAGAAC 6493
Db 6422 TGGAGCAGATCACCGGACATGTCAAAAATGGCTCCATGAGGATTTGTGGGCCAAAAC 6481
Qy 6494 CTGCAGCAACAGTGGCAGGAACTTCCCATCAACGCAATACACACCGGACCTTGAC 6553
Db 6482 CTGCAGCAACAGTGGCATGGAACTTCCCATCAACGCAATACACACCGGACCTTGAC 6541
Qy 6554 ACCCTCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGTGCTGTGAGGAGTACGT 6613
Db 6542 GCCCTCCCGCAGCGCGCACTATTCCAGGGCGCTGTGGGGGTGGCTGCTGAGGAGTACGT 6601
Qy 6614 GGAGTTACGCGTGTGGGGATTTCCACTAGTGAACGGGCATGACCACTGACCAAGTAA 6673
Db 6602 GGAGTTACGCGGTGGGGATTTCCACTAGTGAACGGGCATGACCACTGACCAACGTTGAA 6661

Qy 6674 GTGCCCATGCCAGTTTCGGCCCCCGAATTCTTACGGAGGTGGATGGAGTCGGTTGCA 6733
Db 6662 ATGCCCATGCCAGTTTCAGCCCCCTGAATTTTTCACGGAGGTGGATGGAGTACGGTTGCA 6721
Qy 6734 CAGGTACGCTCCGGCGTGCAAAACCTCTTCTACGGGAGGAGCTCACGTTTCCAGTCCGGCT 6793
Db 6722 CAGGTATGCTCCAGTGTGAAAACCTCTCTACGAGAGGAGTCCGTAATTCAGGTCGGCT 6781
Qy 6794 CAAACATATCTGGTTCGGTTCGAGTCCCATTCGAGCCCCGAAACCGGACGTAAACAGTCT 6853
Db 6782 CAAACAGTACCTGGTTCGGTTCACAGTCCCATTCGAGCCCCGAAACCGGATGTGCACTGCT 6841
Qy 6854 TACTTCCATGCTCACCGATCCCTCCACATTAACAGCAGACGGCTTAAGCTAGGCTGGC 6913
Db 6842 CACTTCCATGCTCACCGACCCCTCTCATATTAACAGCAGACGGCCACGCTAGGCTGGC 6901
Qy 6914 TAGAGGCTCTCCCCCTCTTTTAGCCAGCTCATCAGCTAGCCAGTTGTCTGCGCTCTTTT 6973
Db 6902 CAGGGGCTCTCCCCCTCTTTCGCCAGCTCTTTCAGCTAGCCAGTTGTCTGCGCTCTTTT 6961
Qy 6974 GAAGCGCATGCACTACCCACATGATCTCCCGGAGCGCTGACCTCATCGAGGCCAACCT 7033
Db 6962 GAAGCGCATGCTACTATCCCATCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCT 7021
Qy 7034 CTTTGGCGCAGGAGATGGCGGAAACATCACTCGCTGGAGTCAGAGTAAGGTAGT 7093
Db 7022 CTTTGGCGCAGGAGATGGCGGAAACATCACTCGCTGGAGTCAGAGTAAGGTAGT 7081
Qy 7094 AATTCTGGACTCTTTTCCGAACCGCTTTCACCGGAGGGGATGAGAGGGAGATATCCGCTGC 7153
Db 7082 AATCTGGACTCTTTTCGATCCGATTCGGGGGTGGAGGATGAGAGGGAATATCCGCTCC 7141
Qy 7154 GCGGAGATCTCTGCAAAATCCAGGAATTCCTCCTCAGCGTTGCCCATATGGGACACGCC 7213
Db 7142 GCGGAGATCTCTGCAAAACCCAGGAATTTCCCGCCAGCGTTGCCCATATGGGACACGCC 7201
Qy 7214 GGNCTAATCTCTCACTGCTAGAGTCTCTGGAGGACCGGACTAGCTCCCTCCGCTGCT 7273
Db 7202 GGATTAACAACCTCTCACTGCTAGAGTCTCTGGAAAGGACCGGACTAGCTCCCGCTGGT 7261
Qy 7274 ACACGATGCCATTCGCCACTTCCAAAGGCTCTCCAATACCACTCCACGGAGAAAGAG 7333
Db 7262 ACAGGGTGGCTTTGCCATCTACCAAGGCCCCCCCAATACCACTCCACGGAGGAGAG 7321
Qy 7334 GACGTTGTCTGACAGAAATCCAAATGTCTCTGCTTGGCGGAGCTCGCCACTAAGAC 7393
Db 7322 GACGTTGTCTGACAGAGTCCACCGTGTCTCTGCTTGGCGGAGCTCGCTACTAAGAC 7381
Qy 7394 CTTTCGTTAGTCTCGGATTCGTCGGCGGTTGATAGCGGACCGCGACCGCCCTTCTGACCT 7453
Db 7382 CTTTGGCAGCTCCGGGTCTGTCGGCCGTTGACGCGGACCGCGACTGCGCCCTCCCGATCA 7441
Qy 7454 GGCCTCCGAGCAGCTGCAAAAGGATCCGAGCTTGGAGTCTGCTCTCTTCCATGCCCCCCT 7513
Db 7442 GGCCTCCGAGCAGCGGACAAAGGATCCGAGCTTGGAGTCTGCTCTCTTCCATGCCCCCCT 7501
Qy 7514 TGAAGGGAGCCCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTACCGTGAAGTGA 7573
Db 7502 CGAGGAGAGCCAGGGGACCCCGACCTCAGGACGGGTCTTGGTCTACCGTGAAGTGA 7561
Qy 7574 GGCCTAGGAGTGTGCTGCTGCTCAATGTCTTATACGTGACAGCGGCCCTGTATCAC 7633
Db 7562 AGCTGGTGAAGAGTGTGCTGCTGCTCAATGTCTTATATGAGCAGGTGCTTGTATCAC 7621
Qy 7634 GGCATGGCTCGGAGGAAAGTAAAGTGGCCCATCAACCGTTGAGCAACTCTTTGCTGCG 7693
Db 7622 GGCATGGCTCGGAGGAGAGTAAAGTGGCCCATCAATCCGTTGAGCAACTCTTTGCTGCG 7681
Qy 7694 TCACCAACATGCTGTACGCCCAACATCCCGCAGCGCAAGCCTCCGCGCAGAAAGGT 7753
Db 7682 TCACCAAGTATGCTCTATCCACNACATCTCGCAGCGCAAGTCTGCGGCGAGAGAGGT 7741

QY 7754 CACCTTTGACAGATTGCAAGTCTTGATGATCATTTACCGGACGCTACTCAAGGAGATGAA 7813
Db 7742 CACCTTTGACAGATTGCAAGTCTTGAGACCACTACCGGACGCTCAAGGAGATGAA 7801
QY 7814 GGCAGAGCGCTCCAGATTAAAGGCTTAAGCTTCTATCTATAGAGAGGCGCTCAAGCTGAC 7873
Db 7802 GGCAGAGCGCTCCAGATTAAAGGCTTAGGCTTCTATCTATAGAGAGGCGCTCAAACTGAC 7861
QY 7874 GCCCCCAATTCGCGCAATCCAAATTTGGCTATGGGCAAGGACGTCGCGGACCTATC 7933
Db 7862 GCCCCCAATTCGCGCAATCCAAATTTGGCTATGGGCGAAGGACGTCGCGAGCCTATC 7921
QY 7934 CAGCAGGCGCTTAACCACTCCGCTCCGCTGCGGAGGACTTGTGGAAGACACTGAAAC 7993
Db 7922 CAGCAGGCGCTCAACCACTCCGCTCCGCTGCGGAGGACTTGTGGAAGACACTGAAAC 7981
QY 7994 ACCAAATTGACACCACTATCGGCAAAAGTAGGTTTTCGCTCCAACAGAGAAAGG 8053
Db 7982 ACCAAATTGACACCACTATCGGCAAAAGTAGGTTTTCGCTCCAACAGAGAAAGG 8041
QY 8054 AGCGCGCAAGCGAGCTCGCTTATCGTATTCAGACCTGGAGTTCGTGTATGCGAGAA 8113
Db 8042 AGCGCGCAAGCGAGCTCGCTTATCGTATTCAGACCTGGGCGTACGTGTATGCGAGAA 8101
QY 8114 GATGGCCCTTTACGACGTGGTCTCCACCCTTCCTCAGGCGGTGATGGGCTCCTCATACGG 8173
Db 8102 GATGGCCCTTTACGACGTGGTCTCCACCCTTCCTCAGGCGGTGATGGGCGCTCATACGG 8161
QY 8174 ATTTCAATTAATCCCCCAAGCAGCGGTCGAGTTCCTGTGTAATACCTCGGAAATCAAGAA 8233
Db 8162 ATTCAGTACTCTCTGGGCAAGCGGTCGAGTTCCTGTGTAATACCTCGGAAATCAAGAA 8221
QY 8234 ATGCCCTATGGGCTTCTCATATGACACCGCTGTTTGAATCAAGGCTACGTAGAGTGA 8293
Db 8222 ATGCCCTATGGGCTTCTCATATGACACCGCTGTTTGAATCAAGGCTACGTAGAGTGA 8281
QY 8294 CATTCGTGTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCGGAGGCGACAGAGCG 8353
Db 8282 CATCGTACTGAGGATCAATTTACCAATGTTGTGACTTGGCCCGGAGGCGAGCGAGCG 8341
QY 8354 CATAAGTTCGTCTCAGAGCGGCTTTATCATCGGGGTCCTCTGACTAACTCAAAAGGCA 8413
Db 8342 CATAAGTTCGTCTCAGAGCGGCTTTATGTCGGGGTCCCTGACTAATTCGAAGGGCA 8401
QY 8414 GAACTGCGGTTATCGCCGTCGCGGCAAGTGGGTGCTGACGACTAGCTCGGCTAATAC 8473
Db 8402 GAACTGCGGTTATCGCCGTCGCGGCAAGTGGGTGCTGACGACTAGCTCGGCAACAC 8461
QY 8474 CTTACATGTTACTTGAAGGCCACTGACGCTGTCGAGCTGCAAAAGCTCCAGGACTGCAC 8533
Db 8462 CTTACATGTTACTTGAAGGCCACTGCGGCTGTCGAGCTGCAAGCTCCAGGACTGCAC 8521
QY 8534 GATGCTGCTGAACGGAGACGACCTTGTCTTATCTGTGAAGCGCGGAAACCCAGGAGA 8593
Db 8522 GATGCTGCTGAACGGAGACGACCTTGTCTTATCTGTGAGAGTGGCGGAACCCAGGAGA 8581
QY 8594 TCGCGCGGCTTACGAGCTTCAAGGAGCTATGACTAGGTATTCGCGCCCCCGGGGA 8653
Db 8582 TCGCGCGGCTTACGAGCTTCAAGGAGCTATGACTAGGTATTCGCGCCCCCGGGGA 8641
QY 8654 TCGCGCCCCAAGCAGAAATACGAGCTGAGCTGATATCATGTTCTCCAAATGTGTCAGT 8713
Db 8642 CCGCGCCCCAAGCAGAAATACGAGCTGAGCTGATATCATGTTCTCCAAATGTGTCAGT 8701
QY 8714 CCGCAGCATGATCTGGCAAAAGGCTATACCTCAACCGGTGACCCCAACCCCGCT 8773
Db 8702 CCGCAGCATGATCTGGCAAAAGGCTATACCTCAACCGGTGACCCCAACCCCGCT 8761
QY 8774 TGCAGGCGCTGGTGGGAGACGCTAGACACACTCCAATCAACTCTTGCTAGGCAATAT 8833
Db 8762 CGCAGGCGCTGGTGGGAGACGCTAGACACACTCCAATCAACTCTTGCTAGGCAATAT 8821
QY 8834 CATCATGTATGCGGCCACCCCTATGGGCAAGGATGATTCTGACTACTCTTTTCTCCAT 8893

Db 8822 CATCATGTATGCGGCCACCCCTATGCGGAGGATGATTCTGATGACTCATTTCTCTCTAT 8881
QY 8894 CTTTCTAGCTCAAGAGCAACTTTGAAAAGCCCTCGATTGTGAGATCTACGGGCTTGCTA 8953
Db 8882 CTTTCTAGCTCAAGAGCAACTTTCAAAAAGCCCTCGATTGTGAGATCTACGGGCTTGTA 8941
QY 8954 CTCATTGAGCAGCTTACCTTACCTCAGATCATTTGAACGACTCCATGCTCTTAGCGCAT 9013
Db 8942 CTCATTGAGCAGCTTACCTTACCTCAGATCATTTGAACGACTCCATGCTCTTAGCGCAT 9001
QY 9014 TACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGCTGCTTCATCCCTCAGGAACT 9073
Db 9002 TTCACTCCACAGTTACTCTCCAGGTGAGATCAATAGGCTGCTTCATCCCTCAGGAACT 9061
QY 9074 TGGGTTACACCTTGGGAACCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACT 9133
Db 9062 TGGGTTACCGCTTTGGGAGTCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACT 9121
QY 9134 GTCCAGGGGGAGGCGCGCACTTTGGGAGATACCTTTAACTGGGAGTAAGGAC 9193
Db 9122 GTCCAGGGGGAGGCGTGCACCTTGGGCAAGTACCTTTCAACTGGGAGTAAGGAC 9181
QY 9194 CAAGCTTAAACTCACTCCAATCCCGCGCTCCAGCTGCTGCTGGCTGGTTCGT 9253
Db 9182 CAAGCTTAAACTCACTCCAATCCCGCGCTCCAGCTAGACTTGTCCGCTGGTTCGT 9241
QY 9254 CGCTGTTTACAGCGGGGAGACATATATCACAGCTGTCTCGTGCCTGACCCCGCTGGTT 9313
Db 9242 TGCTGGTTTACAAACGGGGGAGACATATATCACAGCTGTCTCGTGCCTGACCCCGCTGGTT 9301
QY 9314 TCCGTTGCTGCTACTCCCTACTTTCTGTAGGGGTAGGCAATTTACCTGCTCCCCAACCGATG 9373
Db 9302 CATGTTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCAATTTACCTGCTCCCCAACCGATG 9361
QY 9374 AACGGGAGCTTAACCACTCCAGGCTTTAAGGCAATTTCTGTTTTTTTTTTT 9424
Db 9362 AACGGGAGCTTAACCACTCCAGGCAATAGGCCATTCCTTTTTTTTTTTT 9412

RESULT 15

AAQ80498
ID AAQ80498 standard; DNA; 9413 BP.
XX AAQ80498;
XX
DT 16-OCT-2003 (revised)
DT 06-SEP-1995 (first entry)
XX
XX DNA encoding HCV protein cleavable with new serine proteinase.
XX proteinase; serine; cleavage; hepatitis C virus; HCV; ss.
XX Hepatitis C virus; Virus.
XX
XX Key Location/Qualifiers
FT CDS 330..9362
FT /*tag= a
FT /note= "HCV protein"
XX
PN JP06315377-A.
XX
PD 15-NOV-1994.
XX
PF 06-MAY-1993; 93JP-00105666.
XX
PR 06-MAY-1993; 93JP-00105666.
XX
PA (KAEN/) KAENNO K.
PA (SUMQ) SUMITOMO METAL IND LTD.
PA (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.
XX
DR WPI; 1995-032330/05.

Dbb 1862 GAGCCCTGTTAGTGGGGAGCGACCGATCGTTTGGCGGCTCCTACGTAAGCTGGGGGA 1921
Qy 1934 GAAATGAGACAGACGTGATGCTCTCTCAACAAACGCGTCCGCCACAAAGGCAACTGTTCCGG 1993
Db 1922 GAAATGAGACAGACGTGCTGCTACTTAGCAACACGCGCGGCTCAAGGCAACTGTTTGG 1981
Qy 1994 CTGTAATGATGAATAGTACTGAGGTTCACTAAGACGTCGGAGAGTCCCGGTGTAACAT 2053
Db 1982 GTGCACGTGGATGAACAGCACTGGGTTCAACAAACGTCGGGGGCGCTCCGTGCAACAT 2041
Qy 2054 CGGGGGGTCGTAAACCGCACTTGATCTGCCCAACGCACTGCTTCCGGAAGCAACCCCGA 2113
Db 2042 CGGGGGGTCGCAACAAACACCTTGGTCTGCCCAACGCACTGCTTCCGGAAGCAACCCCGA 2101
Qy 2114 GGCTACTTACAAATGTGGTCCGGGCGCTGTTGACACCTAGGTGCTAGTAGACTA 2173
Db 2102 GGCACATTACAAAGTGTGGCTCGGGGCGCTGTTGACACCCAGTGCATGGTTGACTA 2161
Qy 2174 CCATACAGGCTTGGCACTACCCCTGCACTCTCAATTTTTTCCATCTTTAAGGTTAGGAT 2233
Db 2162 CCATACAGGCTCTGGCACTACCCCTGCACTGTTAACTTTTACCGTCTTTAAGGTCAGGAT 2221
Qy 2234 GTATGTGGGGGCTGGAGCACAGGCTCAATGCCCATGCAATTTGGACTCGAGGAGAGCG 2293
Db 2222 GTATGTGGGGGCTGGAGCACAGGCTCAATGCTGCATGCAATTTGGACTCGAGGAGAGCG 2281
Qy 2294 CTGTAACTTGGAGGACAGGATAGTTCAGAACTCAGCCGCTGCTGTCTACAAACAGA 2353
Db 2282 CTGTAACTTGGAGGACAGGATAGTTCAGAACTCAGCCGCTGCTGTCTACAAACAGA 2341
Qy 2354 GTGGCAGATPACTGCCCTGTGCTTTACCAACCTACCGGCTTTATCCACTGCTTTGATCCA 2413
Db 2342 GTGGCAGATPACTGCCCTGTGCTTTACCAACCTACCGGCGCTGCTACCTGGCTTGATCCA 2401
Qy 2414 TCTCCATCAGAACATCGTGGAGCGTGCAATPACTGTACGCTGTAGGCTGAGGCTTGTGCTC 2473
Db 2402 TCTTACCGGAACATCGTGGAGCGTGCAATPACTGTACGCTGTAGGCTGAGGCTTGTGCTC 2461
Qy 2474 CTTTGCMAATCGGAGTACATCTGTTGCTTTTCTTCTCCTGGCAGACGCGCGGT 2533
Db 2462 CTTTGCMAATCGGAGTATATCTGTTGCTTTTCTTCTTCTGGGAGCGCGCGGT 2521
Qy 2534 GTGTGCTCTTGTGGATGATGCTGTATAGCCAGGCTGAGGCGGCTTAGAGAACTT 2593
Db 2522 CTGTGCTCTTGTGGATGATGCTGTATAGCCAGGCTGAGGCACTTAGAGAACTT 2581
Qy 2594 GTGTGCTCAATCGGCGCTCGTGGCGGAGCGCATGTATCTCTCTCTTCTTGTGTT 2653
Db 2582 GGTGTGCTCAATCGGCGCTCTGTGGCGGAGCGCATGTGCTCTCTCTCTCTGTT 2641
Qy 2654 CTTTCTGCGCGGCTGTGTACATTAAGGCGAGGCTGGCTCCTGGGCGGCGTATGCTTTTGA 2713
Db 2642 CTTTCTGCGCGGCTGTGTACATTAAGGCGAGGCTGGCTCCTGGGCGGCGTATGCTTTTGA 2701
Qy 2714 TGGCGTATGGCGGCTGCTCTGCTCTACTGGGCTTACCACACAGGCTTAGGCTTTGGA 2773
Db 2702 TGGCGTATGGCGGCTGCTCTGCTCTTGTGGCTTACCACACAGGCTTAGGCTTTGGA 2761
Qy 2774 CCGGGAGATGGCTGCACTCGTGGGGGCTGCGGTTCTTGTAGTCTGGTATCTTGAACCTT 2833
Db 2762 CCGAGAGATGGCTGCACTCGTGGGAGCGGCTTTTGTAGTCTGGTACTCTTGAACCTT 2821
Qy 2834 GTCCACATACTAACAGTGTCTCTACTAGGCTCATATGGTGTACATCTTTATAC 2893
Db 2822 GTCCACATACTAACAGTGTCTCTCTAGGCTCATATGGTGTACATCTTTATAC 2881
Qy 2894 CAGAGCGAGGCGCACATCAAGTGTGGTCCCGGCTCAAGCTTCCGGGAGGCGCGA 2953
Db 2882 CAGAGCGAGGCGCACATCAAGTGTGGTCCCGGCTCAAGCTTCCGGGAGGCGCGA 2941
Qy 2954 TGCCATCATCTCTCTCACGCTGTGCGGTTCTATCCAGAGTTAAATTTTGAATCACCAGCT 3013
Db 2942 TGCCATCATCTCTCTCACGCTGTGCGGTTCTATCCAGAGCTAAATCTTTTGAATCACCAGCT 3001

Qy 3014 CTTGCTCGCCATATCTCGGCGGCTCATGTGTCTCAGGCTGGCATTAACGAGTGCCTGA 3073
Db 3002 CTTGCTCGCCATATCTCGGCTCGCTCATGTGTCTCAGGCTGGCATTAACGAGTGCCTGA 3061
Qy 3074 CTTGCTCGGCTCAAGGCTCATTTGTCGTCATGTATGTAGTCGGAAGTCCGCGGG 3133
Db 3062 CTTTGTACGCGCTCAGGCGCTCATCCGTCATGTATGTAGTCGGAAGTCCGTCGAGG 3121
Qy 3134 TCATTATGTCCAAATGGTCTTTCATGAAGCTGGGCGCTGACAGGTACGTACGTTTATAA 3193
Db 3122 CCATATGTCCAAATGGCTTTCATGAAGCTGGCGCTGACAGGTACGTACGTTATGA 3181
Qy 3194 CCATCTTACCCACTGCGGGAATCGGCGCACGCGGCGCTACGAGACTTTCGCGTGGCGGT 3253
Db 3182 CCATCTTACTCCAATCGGGGATTTGGCGCACGCGGCGCTACGAGACTTTCGCGTGGCAGT 3241
Qy 3254 AGAGCCGCTGCTCTTCTCCGCCATGGAGACCAAGGTCTACCTTGGGAGCAGACACCGC 3313
Db 3242 AGAGCCGCTGCTCTTCTCTGACATGGAGACTTAAATCTCATACCTTGGGGGCGAGACACCGC 3301
Qy 3314 TGCGTGTGGGACATCATCTTTGGGTCTACCGTCTTCGCGCCGAAAGGGGGAAGGATATT 3373
Db 3302 GCGTGTGGGACATCATCTCGGCTTACCACTTCCGCCCGAAGGGGGAAGGATATT 3361
Qy 3374 TTTGGGACCGGCTGATAGTCTCGAAAGGGCAAGGGTGGCGACTCCTTTCGCCCATCACGCG 3433
Db 3362 TCTAGGACCGGCGGATAGTTTGGAGAGCAGGGTGGCGCTCCTTTCGCCCATCACGCG 3421
Qy 3434 CTACTCCAAACAAAGCGGGGCGTACTTGGTGTGATCATCTAGCTTACAGGCGGGA 3493
Db 3422 CTATTTCCAAACAAAGCGGGGCGTCTTGGCTGTATCATCTAGCTTACAGGCTTACAGGCT 3481
Qy 3494 CAAGAACAGGTCGAAGGGAGGTTCAAGTGTCTTACCGCAACACAACTTTTCTCTGCG 3553
Db 3482 CAAGAACAGGTCGATGGGAGGTTCAAGTGTCTTCCACCGCAACGCAATCTTTCTGCG 3541
Qy 3554 GACCTTGATCAACCGGCTGTCTGGACTGTCTACCATGGCGTGTCTCGAAGACCTTACG 3613
Db 3542 GACCTGTGCAATGGCGTGTGTGGACCGTCTACCATGTGCGGCTCGAAGACCTTACG 3601
Qy 3614 CGGTCCAAAGGTCGAATCAACCAAAATGTAACCAATGTAGACTGTGAGCTTCTGCGGTG 3673
Db 3602 CGGCGCAAGGTCGAATCAACCAAAATGTACCAAAATGTAGACCAAGGCTTCTGCGGTG 3661
Qy 3674 GCAGGCGCGCGGGGCGGCTCCATGACCACTGAGCTGTGAGCTGCGAGCTCGGACTTTA 3733
Db 3662 GCGGCGCGCGCGGGGCGGCTCCATGACACCGGCTCCATGACACCGGCTGCGGAGCTCGGACTTTA 3721
Qy 3734 CTTTGTTCAGAGACATGCTGATGTCTTCCGCTGCGCGGCGGAGCGAGCAGCAGGGGAAG 3793
Db 3722 CTTTGTTCAGAGGACATGCTGATGTCTTCCGCTGCGCGGCGGAGCAGCAGGGGAG 3781
Qy 3794 TCTACTCTCCCGAGGCGGCTCTCTACTGAAAGGCTCTCGGGTGTGCTTCTGCTTGTG 3853
Db 3782 CTTGCTTTCCCGAGGCGGCTCTCTACTGAAAGGCTCTCTCGGCTGAGGCTTCTGCTTGTG 3841
Qy 3854 CTTTGGGGGACGTCGTCGGGCTTCTCGGCTGCTGTGTGCAACCGGGGCTGCGGAA 3913
Db 3842 CTTTGGGGGACGTCGTCGGGCTTCTCGGCTGCTGTGTGCAACCGGGGCTTTCGAA 3901
Qy 3914 GCGGCTGAGCTTCTACACCGGCTTGTGCTTATGAAACTTACCATGCTTCTTCTAC 3973
Db 3902 GCGGCTGAGCTTCTACACCGGCTTGTGCTTATGAAACTTACCATGCTTCTTCTAC 3961
Qy 3974 AGACAACTCAACCCCGGCTGTACCGCAGACATTTCAAGTGGCACTTCTGACGCTCC 4033
Db 3962 AGACAACTCAACCCCGGCTGTACCGCAGGCTTCCGCTTCTTCTAC 4021
Qy 4034 TACTGGCAGGCGCAGAGCACCAAGTGGGCTGCTGTGTGCAACCGGGGCTGCAAGGT 4093
Db 4022 CACTGGCAGGCGCAGAGCACCAAGTGGGCTGCTGTGTGCAACCGGGGCTTCAAGGT 4081

QY 4094 GCTCGTCTGAACCCGCTCCGTTGCCGCCACCTTTAGGGTTGGGGCGTATATATGTCGAAGGC 4153
DB |||||
QY 4082 GCTCGTCTGAACCCGCTCCGTTGCCGCCACCTTTAGGGTTGGGGCGTATATATGTCGAAGGC 4141
DB |||||
QY 4154 ACAGGTATCGACCCCTAAACATCAGAACTGGGTAAAGGACATTTACACGGGGCGGCTCCAT 4213
DB |||||
QY 4142 ACATGCACTGAGCCCTAAACATCAGAACTGGGTAAAGGACATTTACACGGGGCGGCCCCAT 4201
DB |||||
QY 4214 TAGCTACTCCACCTATGCGCAAGTTCTTGCCGACGGTGGCTGTTCTGGGGCGGCTTATGA 4273
DB |||||
QY 4202 CAGTACTCCACCTATGCAAGTTCTTGCCGACGGTGGATGCTCGGGGGCGGCTTATGA 4261
DB |||||
QY 4274 CATCATATATGATGAGTGCACACTCAACTGACTCGACTCAATCTTTGGGCAATCGGCAC 4333
DB |||||
QY 4262 CATCATATATGATGAATGCCACTCAACTGACTCGACTCAATCTTTGGGCAATCGGCAC 4321
DB |||||
QY 4334 AGTCTTGACCAAGCGGAGACGGCTGGAGCGGCTCGTCTGCTCGCCACCGCTACAC 4393
DB |||||
QY 4322 AGTCTTGATCAGGACAGACGGCTGGAGCGGCTCGTCTGCTCGCCACCGCCACGCC 4381
DB |||||
QY 4394 TCCGGATCGGTTACGTTGCCACACCCCAATATCGAGGAATAGGCTGTCCAAACAATGG 4453
DB |||||
QY 4382 TCCGGATCGATCACCGTGCACACCCCAATCGAGGAATGGCTGTCCAAACAATGG 4441
DB |||||
QY 4454 AGAGTCCCCTTCTATGGCAAGCCATCCCATTTGAGGCCATCAAGGGGGGGAGGCACTCT 4513
DB |||||
QY 4442 AGAGTCCCCTTCTATGGCAAGCCATCCCATTTGAGGCCATCAAGGGGGGGAGGCACTCT 4501
DB |||||
QY 4514 CATTTTCTGCAATTTCCAAGGAAGAAATGTGACGAGTCCGCGCAAAAGCTGACAGGCTTCGG 4573
DB |||||
QY 4502 CATCTTCTGCCATTTCCAAGGAAGAGTGTGACGAGTCCGCGCAAAAGCTGACAGGCTTCGG 4561
DB |||||
QY 4574 ACTGAACGCTGTAGCATATTAACGGGGCTTTGATGTGCTCGTCTATACGGCTATCGGAGA 4633
DB |||||
QY 4562 ACTCAATGCTGTAGCGTATTAACCGGGTCTCGATGTGCTCGTCTATACCGATAGCGAGA 4621
DB |||||
QY 4634 CGTGGTTGTGTCGACACAGAGCGCTTAATGACGGGTTTACCGGCGATTTTGACTCAGT 4693
DB |||||
QY 4622 CGTGGTTGTGTCGTCGACACAGAGCGCTTAATGACGGGTTTACCGGCGATTTTGACTCAGT 4681
DB |||||
QY 4694 GATGACTGCAATACATGTGTACCCAGACAGTGTGACTTCACTGCTTGGATCCCACTTCCAC 4753
DB |||||
QY 4682 GATGACTGCAACACATGTGTACCCAGACAGTGTGACTTCACTGCTTGGATCCCACTTCCAC 4741
DB |||||
QY 4754 CATTGAGACGACGACCGTGTCCCAAGACGCGGTGTGCGGTTCGCAACGCGGAGTAGAAC 4813
DB |||||
QY 4742 CATTTGAGACGACGCTGTCCCAAGACGCGGTGTGCGGTGCGGAGCGGCGAGGTAGGAC 4801
DB |||||
QY 4814 TGGCAGGGGTAGGATGGCATCTACAGTTTGTGACTTCAGGAGAAACGGCCCTCGGGCAT 4873
DB |||||
QY 4802 TGGCAGGGGAGGATGGCATCTACAGTTTGTGACTTCAGGAGAAACGGCCCTCAGGCAT 4861
DB |||||
QY 4874 GTTGATTTCTCGGTCGTGTGAGTGTACTAGACGGGCTGTGCTGGTATGAGCTCAC 4933
DB |||||
QY 4862 GTTGACTCTCTCGGTCGTGTGAGTGTATGACGACGCTGCGCTTGGTATGAGCTCAC 4921
DB |||||
QY 4934 GCGGCTCAGACCTCGGTTAGTTGCGGGCTTACCTAAATACACAGGGTTGCGCGTCTG 4993
DB |||||
QY 4922 GCCGCTCAGACCTCGGTTAGTTGCGGGCTTACCTAAATACACAGGGTTGCGCGTCTG 4981
DB |||||
QY 4994 CCAGGACCATCTGGAGTTCTGGGAGAGGCTTTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGC 5053
DB |||||
QY 4982 CCAGGACCATCTAGAGTTCTGGGAGAGGCTTTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGC 5041
DB |||||
QY 5054 CTTCTGTCCAGACTAAACAGGACGAGAGCAACTTTCTTCACTGTGGTGGGATCAAGC 5113
DB |||||
QY 5042 CTTCTGTCCAGACTAAACAGGACGAGAGCAACTTTCTTCACTGTGGTGGGATCAAGC 5101
DB |||||
QY 5114 TACAGTGTGGCGCAGGCTCAAGCTTCACTGTGGGACCAAAATGTGGAAGTGTCT 5173
DB |||||
QY 5102 CACAGTGTGGCGCAGGCTCAGGCTCCACCTCATCGTGGGACCAAAATGTGGAAGTGTCT 5161
DB |||||
QY 5174 CATACGGCTGAACCTACACTGCACGGGCCAACACCCCTGTGTATAGGCTAGGAGCCGT 5233

DB 5162 CATACGGCTAAAGCCACACTGCATGGGCAAGCCCTGCTGTACAGGCTAGGAGCCGT 5221
QY |||||
DB 5234 CCAAAATAGAGTCACTCTCACACACCCCATAACTAAATACATCATGGCAGTATGTCGCG 5293
DB |||||
DB 5222 TCAAAATAGAGTCACTCTCACACACCCCATAAACCAAATACATCATGGCAGTATGTCGCG 5281
QY |||||
DB 5294 TGACCTGGAGGTGCTCACTAGCACCTGGGTGCTGTAGGCGGAGTCTTTGCGAGCTTTGGC 5353
DB |||||
DB 5282 TGACCTGGAGGTGCTCACTAGCACCTGGGTGCTGTAGGCGGAGTCTTTGCGGCTCTGGC 5341
QY |||||
QY 5354 CGCATACTCCCTGACGACAGGAGTGTGCTCACTTTGGGCGAGGATCATCTTTGTCGGGAA 5413
DB |||||
DB 5342 CGGCTACTGCTGACACAGGAGAGTGTGCTCACTTTGGGCGAGGATCATCTTTGTCGGGAG 5401
QY |||||
DB 5414 GCAGCTGTGCTTTCCCGACAGGGAAGTCTCTACACGAGGATTCGATGAGATGGAAGAGTG 5473
DB |||||
DB 5402 GCCAGCTGTATTTCCCGACAGGGAAGTCTCTACACGAGGATTCGATGAGATGGAAGAGTG 5461
QY |||||
DB 5474 TGCTCTCAAACTTCTTTAATCGAGCAGGGAATGCACTGCGCGGAGCAATTTCAAGCAAAA 5533
DB |||||
DB 5462 TGCTCTCAAACTTCTTTAATCGAGCAGGGAATGCACTGCGCGGAGCAATTTCAAGCAAAA 5521
QY |||||
DB 5534 GCGCTCGGGTGTGTCAAACCGCCACCAAGCAAGCGGAGGCTGCTCCCGTGTGGA 5593
DB |||||
DB 5522 GCGCTCGGATGCTGTCAAACAGCCACCAAGCAAGCGGAGGCTGCTCCCGTGTGGA 5581
QY |||||
DB 5594 GTCCAAGTGGCGAGCCCTTGAGACCTTCTGGGCGAAGACACATCTGGAATTTTCATCAGCGG 5653
DB |||||
DB 5582 GTCCAAGTGGCGAGCCCTTGAGGTCTTCTGGGCGAAGACACATCTGGAATTTTCATCAGCGG 5641
QY |||||
DB 5654 AATACAGTACTAGCAGGCTTATCCACTCTGCTCGGAACCCCGCGATAGCATCAATGAT 5713
DB |||||
DB 5642 GATACAGTACTTGGCAGGCTATCCACTCTGCTCGGAACCCCGCGATAGCATCAATGAT 5701
QY |||||
DB 5714 GGCAATTTACAGTCTATACAGCCGCTTACACCCCAAAACACCCCTCTCTGTTAATCAT 5773
DB |||||
DB 5702 GGCTTTTACAGCTCTATACAGCCGCTTACACCCCAAAATACCCCTCTCTGTTTAAATCAT 5761
QY |||||
DB 5774 CTTGGGGGATGGTGGTGGCTGCCAACTGCTCTCCAGCGCTGCGTCACTTTTCGTTGGG 5833
DB |||||
DB 5762 CTTGGGGGATGGTGGTGGCTGCCAACTGCTCTCCCGCGAGGCTGCTTCCGCTTTTCGTTGG 5821
QY |||||
DB 5834 CGCGGCTCGCGAGCGGCTGTTGGCAGCATAGGCTTTGGGAAGTGTGCTGCTGAGCAT 5893
DB |||||
DB 5822 CGCGGCTCGCGAGCGGCTGTTGGCAGCATAGGCTTTGGGAAGTGTGCTGCTGAGCAT 5881
QY |||||
DB 5894 CTTGGCGGCTATGGGCGAGGCTGAGCGGCGCATCTGCTGCGCTTTAAGGTTCATGAGCGG 5953
DB |||||
DB 5882 TCTGGCGGCTATGGGCGGCGGCTGAGCGGCGCATCTGCTGCGCTTTAAGGTTCATGAGCGG 5941
QY |||||
DB 5954 CGAGTCCCTTCCACCGAGGACCTGGTCAACTTACTCTCTGCGCATCTCTCTCTGCTGTC 6013
DB |||||
DB 5942 CGAGTCCCTTCCACCGAGGATCTGGTTAATTTACTCTCTGCGCATCTCTCTCTGCTGTC 6001
QY |||||
DB 6014 CTTGGTCTGCGGGTGTGTCGCGACCAATACTGCTGCGACAGTGGGCGCGGAGAGG 6073
DB |||||
DB 6002 CTTGGTCTGCGGGTGTGTCGCGACCAATACTGCTGCGGACAGTGGGCGCGGAGAGG 6061
QY |||||
DB 6074 GGCTGTGCGGATGAGGACCGGCTGATAGGCTTCTGCTTGGGGGTAAACACGCTCTCTCCC 6133
DB |||||
DB 6062 GGCTGTGCGGATGAGGACCGGCTGATAGGCTTCTGCTTGGGGGTAAACACGCTCTCCC 6121
QY |||||
DB 6134 TACGCACTATGCTGAGAGGACGCTGACAGCAGTGTCTCACTCAGATCTCTCTAGCCT 6193
DB |||||
DB 6122 CACGCACTATGTCGCGAGGAGCGCGCGCGGCTGTACTCAGATCTCTCTCAGCCT 6181
QY |||||
DB 6194 TACCATCACTCAACTGCTGAACGGGCTCCACAGTGGATTAATGAGGACTGCTCTACGCG 6253
DB |||||
DB 6182 TACCATCACTCAGTTGCTGGAAGGCTTCACTCAGTGGATTAATGAGGACTGCTCTCAGCG 6241
QY |||||
DB 6254 ATGCTCCGGCTGCTGGCTGAAGGATGTTGGGATGGATATGACAGGTTGACTGACTT 6313

Db 6242 TTGTTCCGGCTCGTGGCTAAAGGATGTTTGGGACTGGATATGCA CGGTGTTGAGTGACTT 6301
QY 6314 CAAGACCTGGCTCCAGTCCAAACTCCTGCGCGGTTTACCGGAGTCCCTTTCTCTGTCATG 6373
Db 6302 CAGACTTGGCTCCAGTCCNAGCTCTGCGCGGTTTACCGGAGTCCCTTTCTCTGTCATG 6361
QY 6374 CCAA CGCGGTTACAAGGAGTCTGCGGGGAGCGGCATCATGCAAAACACCTGCCCATG 6433
Db 6362 CCAA CGCGGTTACAAGGAGTCTGCGGGGAGTGGCATCATGCAAAACACCTGCCCATG 6421
QY 6434 CGGAGCAGATCGCGGAGCATGTCAAAAACCGTTCCATGAGGATCTGAGGCTAGAAC 6493
Db 6422 TGGAGCAGATGACCGGAGCATGTCAAAAATGGCTCCATGAGGATGTTGGGCCAAAAC 6481
QY 6494 CTGAGCAACAGTGGCAGGAAAGTTCCCATCAACGATACACACCGGAGCTTGGAC 6553
Db 6482 CTGAGCAACAGTGGCAGTGGAAACATTTCCCATCAACGATACACACCGGCGCCGTGCAC 6541
QY 6554 ACCCTCCCGCGGCCCAACTATTTCAGGGCGCTATGGCGGTTGGCTGTGAGGAGTACGT 6613
Db 6542 GCCCTCCCGCAGCGCGAACTATTTCAGGGCGCTGTGGCGGTTGGCTGTGAGGAGTACGT 6601
QY 6614 GGAGTTAGCGTGTGGGGATTTCACTAGTGAAGGCGATGACCACTGACAAACGTAAA 6673
Db 6602 GGAGTTAGCGCGGTGGGGATTTCACTAGTGAAGGCGATGACCACTGACAAACGTGAA 6661
QY 6674 GTGCCATGCCAGGTTCCGCGCCCGAAATTTCTTACGGAGTGGATGGAGTSCGTTGCA 6733
Db 6662 ATGCCATGCCAGGTTCCAGCCCTCGAATTTTTCAGGAGTGGATGGAGTACGTTGCA 6721
QY 6734 CAGGTACGCTCGCGGTGCAAACTCTTCTACGGGAGGACGTACGTTTCCAGTCCGGCT 6793
Db 6722 CAGGTATGCTCAGTGTGCAAACTCTCTACGAGAGGAGTCTGTAATTCAGTCCGGCT 6781
QY 6794 CAAACAACTATTGGTTCGGTTCGAGTCCAGTCCATGCGAGCCCGAACCGAGCTAACAGTGT 6853
Db 6782 CAAACAGTACCTGGTTCGGTTCACAGTCCCATGTGAGCCCGAACCGAGTGTGGCAGTGT 6841
QY 6854 TACTTCCATGCTCACGATCCCTCCACATTAACAGAGAGCGCTAAGCGTAGGCTGGC 6913
Db 6842 CACTTCCATGCTCACGACCTCTCATATTAACAGAGAGCGGCAAGCGTAGGCTGGC 6901
QY 6914 TAGAGGTTCTCCCTCTTTAGCCAGTCTCATGAGTACGAGTGTGTGCGCTTCTTT 6973
Db 6902 CAGGGGTTCTCCCTCTCTTCCGAGCTCTTACGAGTACGAGTGTGTGCGCTTCTTT 6961
QY 6974 GAAAGGACATGCACTACCACTGATCTCCCGAGCGTACCTCATCGAGGCGCAACT 7033
Db 6962 GAAAGGACATGCACTACCACTGATCTCCCGAGCGTACCTCATCGAGGCGCAACT 7021
QY 7034 CTTGTGGCGGAGGAGATGGCGGAAACATCACTCCGTTGGAGTCAGAGNATAAGGTAGT 7093
Db 7022 CTTGTGGCGGAGGAGATGGCGGAAACATCACTCCGTTGGAGTCAGAGNATAAGGTAGT 7081
QY 7094 AATTCTGGACTCTTTGCAACCGCTTTCAGCGGAGGAGATGAGAGGAGATATCGTGGC 7153
Db 7082 AATCTGGACTCTTTGCACTCCGATTCGCGGAGTGGAGTGGAGGAAATATCGTCCC 7141
QY 7154 GCGGAGATCTCTGCAAAATCCAGAAAGTTCCCTCAGCGTTGCCATATGGGACGCCC 7213
Db 7142 GCGGAGATCTCTGCAAAACCCAGAAAGTTCCCTCCAGCGTTGCCATATGGGACGCCC 7201
QY 7214 GGACTACATCTCCACTGCTAGTCTCTGGAGACCGGACTACGCTCCCTCCGTTGGT 7273
Db 7202 GGAATTACAACTCTCCACTGCTAGTCTCTGGAGGATCCGGACTACGTTCCCTCCGTTGGT 7261
QY 7274 ACACGGATGCCCATTTGCCACTTACCAAGGCTCTCTCAATACCCTTCCACGAGAGAAAG 7333
Db 7262 ACACGGGTGCCCTTTGGCATCTTACCAAGGCCCTCCCATACCCTTCCACGAGGAGAG 7321
QY 7334 GACGTTGTCTGACAGAAATCCAAATGTGTCTGCTTGGCGGAGTCCGCACTAAGAC 7393
Db 7322 GACGTTGTCTGACAGAGTCCACCGCTCTTCTGCTTGGCGGAGTCTGCTACTAAGAC 7381

QY 7394 CTTGGTAGTCTCCGATCGTCCGCGGTTGATAGCGGACCGGCGACCGCCTTCTTGACCT 7453
Db 7382 CTTTGGCAGCTCCGGTCTGTCGGCGTGTGACAGCGGACAGGCGACTGGCCCTCCGATCA 7441
QY 7454 GGCCTCCGACGAGGTGACAAAGATCCGAGTTGAGTCTGTACTCTCTCATGCCCCCT 7513
Db 7442 GGCCTCCGACGAGCGGACAAAGATCCGAGTTGAGTCTGTACTCTCTCATGCCCCCT 7501
QY 7514 TGAAGGGAGCGGGGACCCGATCTCAGCGACGGGCTTTGGTCTACGGTGAAGGAG 7573
Db 7502 CGAGGAGAGCGAGGGACCCCGACCTCAGCGACGGGCTTTGGTCTACGGTGAAGGAG 7561
QY 7574 GGTAGTGAAGATGCTGCTGCTCAATGTCTTATACGTGGACAGCGCCCTCATAC 7633
Db 7562 AGCTGTGAGAGCTGCTGCTGCTCAATGTCTTATACATGAGCAGGTGCTTGTATC 7621
QY 7634 GCCATGCTCGGAGGAAAGTAAAGTCCCATCAACCCGTTGAGCAACTCTTTCTGCG 7693
Db 7622 GCCATGCTCGGAGGAGAGCAAGTTGCCCATCAATCCGTTGAGCAACTCTTTCTGCG 7681
QY 7694 TCACCAACATGCTTACGCGACCAACATCCGCGAGCGCAAGCCTCCGCGAGAGAGT 7753
Db 7682 TCACCAACATGCTTACTCCCAACATCTCGCAGCGCAAGTCTCGCGCAGAGAGT 7741
QY 7754 CACCTTTGACAGATTGCAAGTCTTGGATGATCATTTACCGGACGTACTCAAGGAGATGA 7813
Db 7742 CACCTTTGACAGATGCAAGTCTTGGACGACCACTACCGGAGCGTCTCAGGAGATGA 7801
QY 7814 GCGAAGCGCTCCAGTTAAGCTTCTATCTATAGAGAGGCTCGCAAGCTGAC 7873
Db 7802 GCGAAGCGCTCCAGTTAAGCTTCTATCTATAGAGAGGCTCGCAAGCTGAC 7861
QY 7874 GCCCCCATTTCCGCAAAATCCAAATTTGGCTATGGGCAAGAGCGTCCGGAACCTATC 7933
Db 7862 GCCCCCATTTCCGCAAAATCCAAATTTGGCTATGGGCAAGAGCGTCCGAGGCTATC 7921
QY 7934 CAGCAGGCGCTTAAACACATCCGCTCGTGTGGGAGACTTGTCTGGAAGACACTGAAAC 7993
Db 7922 CAGCAGGCGCTTAAACACATCCGCTCGTGTGGGAGACTTGTCTGGAAGACACTGAAAC 7981
QY 7994 ACCAATTGACACCACTATCGCAAAAGTGGGTTTTCTGCGTCCAAACAGAGAGG 8053
Db 7982 ACCAATTGACACCACTATCGCAAAAGTGGGTTTTCTGCGTCCAAACAGAGAGG 8041
QY 8054 AGCGCGAAGCGCTCGCTTATCGTATTCAGACCTGGAGTTCGTGTATGCGAGAA 8113
Db 8042 AGCGCGAAGCGCTCGCTTATCGTATTCAGACCTGGGAGTTCGTGTATGCGAGAA 8101
QY 8114 GATGCGCTTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGGGCTCTCATACGG 8173
Db 8102 GATGCGCTTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGGGCTCTCATACGG 8161
QY 8174 ATTTCAATPACTCCCGCAAGCAGCGGTCTGAGTTCTGTTGAATACCTGGAATCAAGAA 8233
Db 8162 ATTTCAATPACTCCCGCAAGCAGCGGTCTGAGTTCTGTTGAATACCTGGAATCAAGAA 8221
QY 8234 ATGCGCTATGGCTTCTCATATGACACCGCTGTTTGACTCAACCGGTCTCATGAGTGA 8293
Db 8222 ATGCGCTATGGCTTCTCATATGACACCGCTGTTTGACTCAACCGGTCTCATGAGTGA 8281
QY 8294 CATTCGCTTTCAGGAGTCAATTTTACCAATGTTGAGTTCGCGCCCGGCGGAGAGCGG 8353
Db 8282 CATTCGCTTTCAGGAGTCAATTTTACCAATGTTGAGTTCGCGCCCGGAGAGCGGAG 8341
QY 8354 CATAAGGTCTGCTCAGAGCGGCTTTTATCATCGGGGTCTCCCTGACTAATCAAAAGGCA 8413
Db 8342 CATAAGGTCTGCTCAGAGCGGCTTTTATCATCGGGGTCTCCCTGACTAATTCGAAGGCA 8401
QY 8414 GAACTGCGGTTATCGCGGTCGCGGCAAGTGGCGTGTCTGAGCACTAGTCTGGGTAAATC 8473
Db 8402 GAACTGCGGTTATCGCGGTCGCGGCAAGTGGCGTGTCTGAGCACTAGTCTGGGTAAATC 8461

Qy	8474	CTCATATGTTACTTGAAGCCACTGAGAGCTGTGAGCTGCAAAAGCTCCAGSACTGCAC	8533
Db	8462	CTTCATATGTTACTTGAAGCCACTGCGGCTGTGAGCTGCAAAAGCTCCAGSACTGCAC	8521
Qy	8534	GATGCTCGTGAACGGAGAGACCTTGTGTTATCTGTGAAAGCGCGGGAACCCAGGAGGA	8593
Db	8522	GATGCTCGTGAACGGAGAGACCTTGTGTTATCTGTGAGAGTGCAGGAAACCCAGGAGGA	8581
Qy	8594	TGCGGGGGCCCTACGAGCCCTTACGAGGCTATGACTAGGTATTCGCGCCCCCCCCGGGGA	8653
Db	8582	TGCGGGGGCCCTACGAGCCCTTACGAGGCTATGACTAGGTATTCGCGCCCCCCCCGGGGA	8641
Qy	8654	TCGCGCCCAACCAAGATACAGACCTGAGAGCTGATAACATCATGTTCTCCAAATGTTCACT	8713
Db	8642	CCGCGCCCAACCAAGATACAGACTTGGAGCTGATAACGTCATGCTCTCCAAATGTTGCGT	8701
Qy	8714	CGCGCAGATGATCTGGCAAAAGGTATACTACCTCACCCGTGACCCACACCCGCCCT	8773
Db	8702	CGCGCAGATGATCTGGCAAAAGGTATACTACCTCACCCGTGACCCACACCCGCCCT	8761
Qy	8774	TGCACGGGCTGCGTGGGAGACAGCTAGACACACTTCCAAATCAACTCTTGGCTAGGCAATAT	8833
Db	8762	CGCACGGGCTGCGTGGGAGACAGTTAGACACACTCCAGTCAACTCTTGGCTAGGCAATAT	8821
Qy	8834	CATCATGTATGCGCCCAACCTATGCGCAAGGATGATTTCTGATGACTCACTTTTCTCCAT	8893
Db	8822	CATCATGTATGCGCCCAACCTATGCGCCAGGATGATTTCTGATGACTCACTTTCTCTAT	8881
Qy	8894	CTTCTAGCTCAGAGCAACTTGAAGAGCCCTGGATTGTCAGATCTACGGGGCTTGCTA	8953
Db	8882	CTTCTAGCTCAGAGCAACTTGAAGAGCCCTGGATTGTCAGATCTACGGGGCTTGTTA	8941
Qy	8954	CTCATTTAGCCACTTGACCTACCTCAGATCAATTGAAACGACTCCATGGTCTTAGCGCAT	9013
Db	8942	CTCATTTAGCCACTTGACCTACCTCAGATCAATTGAAACGACTCCATGGTCTTAGCGCAT	9001
Qy	9014	TACATCCAAGTTACTCTCAGGTAGATCAATAGGGTGGCTTCATGCTCAGGAAACT	9073
Db	9002	TTCACTCCAAGTTACTCTCAGGTAGATCAATAGGGTGGCTTCATGCTCAGGAAACT	9061
Qy	9074	TGGGGTACCACCTTGGCAACCTGGAGACATCGGGCCAGAGTGTCCGGCTAAGCTACT	9133
Db	9062	TGGGGTACCACCTTGGCAACCTGGAGACATCGGGCCAGAGTGTCCGGCTAAGCTACT	9121
Qy	9134	GTCCAGGGGGAGGCGCGCACTTGTGGCAGATACCTCTTTAACTGGGCGAGTAAGGAC	9193
Db	9122	GTCCAGGGGGAGGCGCGCACTTGTGGCAGATACCTCTTTAACTGGGCGAGTAAGGAC	9181
Qy	9194	CAAGCTTAAACTCACTCCAATCCCGCGCGGTCCAGCTGGACTTGTCTGGCTGGTTGCT	9253
Db	9182	CAAGCTTAAACTCACTCCAATCCCGCGCGGTCCAGCTGGACTTGTCTGGCTGGTTGCT	9241
Qy	9254	CGCTGGTTACAGCGGGGAGACATATATACAGGCTGTCTGTCGCCGACCCCGCTGGTT	9313
Db	9242	TGCTGGTTACAGCGGGGAGACATATATACAGGCTGTCTGTCGCCGACCCCGCTGGTT	9301
Qy	9314	TCGCTTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCAATTTACCTGCTCCCAACCGATG	9373
Db	9302	CATGTTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCAATTTACCTGCTCCCAACCGATG	9361
Qy	9374	AACGGGAGCTAACCACTCCAGGCTTAAGCCATTTTCTGTTTTTTTTTTT 9424	
Db	9362	AACGGGAGCTAACCACTCCAGGCTTAAGCCATTTTCTGTTTTTTTTTTT 9412	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 08:34:14 ; Search time 291 Seconds
(without alignments)
4321.877 Million cell updates/sec

Title: US-09-662-454-3
Perfect score: 16009
Sequence: 1 MSTNPKPQRTKNTNRPRQ.....FPLCLLLSVGVGYLLPNR 3010

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubppa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15323	95.7	3010	4	US-10-467-000-1
2	15064	94.1	3010	4	US-10-333-449A-34
3	14801.5	92.5	2985	4	US-10-259-275-40
4	14801.5	92.5	2985	6	US-11-006-313-40
5	14162.5	88.5	3011	3	US-09-742-659-4
6	14143.5	88.3	3011	3	US-09-891-894-3
7	14143.5	88.3	3011	4	US-10-184-150-3
8	14143.5	88.3	3011	4	US-10-328-997-3
9	14143.5	88.3	3012	3	US-09-238-076-2
10	14143.5	88.3	3012	3	US-09-995-937-2
11	14143.5	88.3	3012	3	US-09-917-563-2
12	14131.5	88.3	3011	4	US-10-296-734-406
13	14130.5	88.3	3011	6	US-11-126-662-1
14	14124.5	88.2	3011	3	US-09-952-572-9
15	14124.5	88.2	3011	3	US-09-747-419-20
16	14124.5	88.2	3011	4	US-10-259-275-20
17	14124.5	88.2	3011	4	US-10-189-359-14
18	14124.5	88.2	3011	6	US-11-006-313-20
19	14119.5	88.2	3011	3	US-09-916-359-2
20	14119.5	88.2	3011	4	US-10-445-724-2
21	14090.5	88.0	3011	3	US-09-238-076-20
22	14090.5	88.0	3011	3	US-09-995-937-20
23	14090.5	88.0	3011	3	US-09-917-563-20
24	13999.5	87.4	3011	4	US-10-232-643-6
25	13935.5	87.0	3011	3	US-09-929-955-1
26	13935.5	87.0	3011	4	US-10-104-966-1
27	13935.5	87.0	3011	4	US-10-719-619-1

28 13935.5 87.0 3011 5 US-10-817-591-1 Sequence 1, Appli
29 13442.5 84.0 2894 3 US-09-941-611-23 Sequence 23, Appl
30 13442.5 84.0 2894 4 US-10-044-995-23 Sequence 23, Appl
31 13442.5 84.0 2894 5 US-10-822-871-23 Sequence 23, Appl
32 11614.5 72.5 2940 4 US-10-226-629A-13 Sequence 13, Appl
33 11145 69.6 2201 4 US-10-085-476-2 Sequence 2, Appli
34 11127 69.5 2201 4 US-10-029-907-3 Sequence 3, Appli
35 11127 69.5 2201 4 US-10-309-561-3 Sequence 3, Appli
36 11127 69.5 2201 4 US-10-789-355-3 Sequence 3, Appli
37 11127 69.5 2201 4 US-10-886-835-3 Sequence 3, Appli
38 10152 63.4 1985 5 US-10-492-178-6 Sequence 6, Appli
39 10131 63.3 1985 5 US-10-492-178-6 Sequence 1, Appli
40 10106 63.1 1985 4 US-10-639-150-2 Sequence 2, Appli
41 10100 63.1 1985 4 US-10-259-275-42 Sequence 42, Appl
42 10100 63.1 1985 6 US-11-006-313-42 Sequence 42, Appl
43 8272.5 51.7 1892 5 US-10-612-884-6 Sequence 2, Appli
44 7143 44.6 1736 4 US-10-328-127-2 Sequence 2, Appli
45 7143 44.6 1736 4 US-10-328-206-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-467-000-1
; Sequence 1, Application US/10467000
; Publication No. US20040067486A1
; GENERAL INFORMATION:
; APPLICANT: De Francesco, Raffaele
; APPLICANT: Migliaccio, Giovanni
; APPLICANT: Pacnessa, Giovanni
; TITLE OF INVENTION: HEPATITIS C VIRUS REPLICONS AND REPLICON
; FILE OF INVENTION: ENHANCED CELLS
; FILE REFERENCE: ITR0003P
; CURRENT APPLICATION NUMBER: US/10/467,000
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: PCT/EP02/00526
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/263,479
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Con 1 HCV isolate nucleic acid
US-10-467-000-1

Query Match	95.7%	Score 15323	DB 4	Length 3010
Best Local Similarity	94.8%	Pred. No. 0		
Matches 2852	Conservative	76	Mismatches	82
Indels	0	Gaps	0	
Qy	1	MSTNPKPQRTKNTNRPRQDPVGGGOI	VGGVYLLPRRGPRLGVRATRKASERSQPRG	60
Db	1	MSTNPKPQRTKNTNRPRQDPVGGGOI	VGGVYLLPRRGPRLGVRATRKATRSQPRG	60
Qy	61	RRQIPKARRPEGRWAQPGYPWPLYNEGLWAGWLLSPRGS	PSWGTPDPRRSRLG	120
Db	61	RRQIPKARQEGRAWAQPYPWPLYNEGLWAGWLLSPRGS	PSWGTPDPRRSRLG	120
Qy	121	KVIDTLTCGADLMGYIPLVGAPLGGARALAHGVRLVEDG	VNYATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCGADLMGYIPLVGAPLGGARALAHGVRLVEDG	VNYATGNLPGCSFSIFLLA	180
Qy	181	LLSLCTTPASAYEVNRVSGIVHVNDSCNSISIVVEADVM	HTPCVPCVQVGNSSRCWV	240
Db	181	LLSLCTTPASAYEVNRVSGIVHVNDSCNSISIVVEADVM	HTPCVPCVQVGNSSRCWV	240
Qy	241	ALTPTLAARNASVPTTIRRHVDLLVGTAAFC	SAMYVGDLCGSIPLVSQLFTFSRRHET	300
Db	241	ALTPTLAARNASVPTTIRRHVDLLVGTAAFC	SAMYVGDLCGSIPLVSQLFTFSRRHET	300
Qy	301	VQDCNCSIYPGHVSGHRMADMMNWSPTTALVWSQLLR	IPQAVYDMVAGAHWGVLGLA	360

QY	2521	GYGAKDVRNLSRAVNHRSVWEDLLEDTEPIDTTIMAKSEVFCVQPEKGRKPARLIV	2580
DB	2521	GYGAKDVRNLSRAVNHRSVWEDLLEDTEPIDTTIMAKSEVFCVQPEKGRKPARLIV	2580
QY	2581	FPDLGVRVCEKMAIYDVVSTLPOAVNGSSYGFQYSPQKRVFELVNTWKKCPMGFSYDT	2640
DB	2581	FPDLGVRVCEKMAIYDVVSTLPOAVNGSSYGFQYSPQKRVFELVNTWKKCPMGFSYDT	2640
QY	2641	RCFDSVTTESDIRVESIYOCCLDAPARQAIRSLTERLYIGGPLTNSKGNGCYRRARA	2700
DB	2641	RCFDSVTTESDIRVESIYOCCLDAPARQAIRSLTERLYIGGPLTNSKGNGCYRRARA	2700
QY	2701	SGVLTTSCGNLTLCYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFT	2760
DB	2701	SGVLTTSCGNLTLCYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFT	2760
QY	2761	AMTRYAPGDPDPQPEYDLELITSCSNVSVVAHDAGKRVYLYTRDPTTPLARAANETAR	2820
DB	2761	AMTRYAPGDPDPQPEYDLELITSCSNVSVVAHDAGKRVYLYTRDPTTPLARAANETAR	2820
QY	2821	HTPINSWLGNIIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPO	2880
DB	2821	HTPINSWLGNIIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPO	2880
QY	2881	IIEHLGLSFTLHSHSYSPGGINRVASCLRKLGVPPLRTWHRARSVRAKLLSQGGAATC	2940
DB	2881	IIEHLGLSFTLHSHSYSPGGINRVASCLRKLGVPPLRTWHRARSVRAKLLSQGGAATC	2940
QY	2941	GRYLFNMAVTKLKLTPIPAASQDLDSGMFVAGYSGGDIYHLSRARPRWFPICLLLSV	3000
DB	2941	GRYLFNMAVTKLKLTPIPAASQDLDSGMFVAGYSGGDIYHLSRARPRWFPICLLLSV	3000
QY	3001	GVGIYLLPNR 3010	
DB	3001	GVGIYLLPNR 3010	
RESULT 2			
US-10-333-449A-34			
; Sequence 34, Application US/10333449A			
; Publication No. US20040137424A1			
; GENERAL INFORMATION:			
; APPLICANT: Tan, Yin Hwee			
; APPLICANT: Lim, Siew Pheng			
; APPLICANT: Lim, Seng Gee			
; APPLICANT: Hong, Wan Jin			
; TITLE OF INVENTION: NUCLEIC ACIDS AND METHODS FOR DETECTING VIRAL INFECTION,			
; TITLE OF INVENTION: UNCOVERING ANTI-VIRAL DRUG CANDIDATES AND DETERMINING DRUG			
; TITLE OF INVENTION: RESISTANCE OF VIRAL ISOLATES			
; FILE REFERENCE: 01/22137			
; CURRENT APPLICATION NUMBER: US/10/333.449A			
; CURRENT FILING DATE: 2003-01-21			
; NUMBER OF SEQ ID NOS: 34			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 34			
; LENGTH: 3010			
; TYPE: PRT			
; ORGANISM: Hepatitis C virus			
US-10-333-449A-34			
Query Match 94.1%; Score 15064; DB 4; Length 3010;			
Best Local Similarity 93.2%; Pred. No. 0;			
Matches 2805; Conservative 86; Mismatches 119; Indels 0; Gaps 0;			
QY	1	MSTNPKPQKTKRNTNRBPQDKFPGGQIVGVVLLPRRGPRLGVRATRKASERSQPRG	60
DB	1	MSTIPKQKTKRNTNRBPQDKFPGGQIVGVVLLPRRGPRLGVRATRKASERSQPRG	60
QY	61	PROPIPKARRPRGAWAOPGYWPPLYGNEGGLWAGWLLSPGRSRPSWGPTDPRRRSRNLG	120
DB	61	RQPIPKARRPRGAWAOPGYWPPLYGNEGGLWAGWLLSPGRSRPSWGPTDPRRRSRNLG	120

QY	121	KVIDTLTCGFPADLMGYIPLVGAPLGGAAALAHGVRVLEDDGVNYATGNLPGCSFSFIPLA	180
DB	121	KVIDTLTCGFPADLMGYIPLVGAPLGGAAALAHGVRVLEDDGVNYATGNLPGCSFSFIPLA	180
QY	181	LLSCLTIPASAYEVNRVSGIYHVTNDCSNSSIVYEAADIMHTPGCPVCQEGNSRCRW	240
DB	181	LLSCLTIPASAYEVNRVSGIYHVTNDCSNSSIVYEAADIMHTPGCPVCQEGNSRCRW	240
QY	241	ALPTPLAARNAASVPTTTRRHVDLLVGTAAAFCSAMVGDLCGSIPLVSOLFTHSPRHET	300
DB	241	ALPTPLAARNAASVPTTTRRHVDLLVGTAAAFCSAMVGDLCGSIPLVSOLFTHSPRHET	300
QY	301	VODCNCSIPVGHVSHRMAWMMWNSPTTALVVSOLLRIPOAVDMVAGAHGWGLAGLA	360
DB	301	VODCNCSIPVGHVSHRMAWMMWNSPTTALVVSOLLRIPOAVDMVAGAHGWGLAGLA	360
QY	361	YYSWGNWAKVLIIVALLFAGVDGTHHTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW	420
DB	361	YYSWGNWAKVLIIVALLFAGVDGTHHTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW	420
QY	421	HINRTALNCVDSLQGTFFAALFYAHKFNSSGCPERMASCRPIDWPAQCGPITYTKENSS	480
DB	421	HINRTALNCVDSLQGTFFAALFYAHKFNSSGCPERMASCRPIDWPAQCGPITYTKENSS	480
QY	481	DQRPYCHWYAPRPGVVPASOYCGPVYCTPTSPVVVGTTRDSGVPTYWGENETDVMNLN	540
DB	481	DQRPYCHWYAPRPGVVPASOYCGPVYCTPTSPVVVGTTRDSGVPTYWGENETDVMNLN	540
QY	541	NTRPPGNGFGCTWMTNSTGFTKCGGPPCNIGVGNRTLIPTDCFRKHPEATYTKGSG	600
DB	541	NTRPPGNGFGCTWMTNSTGFTKCGGPPCNIGVGNRTLIPTDCFRKHPEATYTKGSG	600
QY	601	PWLTPTCLVDYPIYRLWHYPCTLFNFSIFKVRMYGVGVEHRLNAAACNWTGRGRCNLEDRS	660
DB	601	PWLTPTCLVDYPIYRLWHYPCTLFNFSIFKVRMYGVGVEHRLNAAACNWTGRGRCNLEDRS	660
QY	661	ELSPLLSTTTEWQILPCAFITLIPALSTGLIHLHQNIVDVQYLVXGVSFAFKWYVL	720
DB	661	ELSPLLSTTTEWQILPCAFITLIPALSTGLIHLHQNIVDVQYLVXGVSFAFKWYVL	720
QY	721	LLFLLADARVCAACLMWMLLIAQEAALLENLVNNAASVAGAHGILSFLVFFCAAWIKG	780
DB	721	LLFLLADARVCAACLMWMLLIAQEAALLENLVNNAASVAGAHGILSFLVFFCAAWIKG	780
QY	781	RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCGGAVLVGLVFLTSPYVKFLT	840
DB	781	RLAPGAAYAFYGVWPLLLLLALPPRAYALDREMAASCGGAVLVGLVFLTSPYVKFLT	840
QY	841	RLIWLQYFTRABAHQVWVPLNVGRGRDAIILLTCAVHPELIIDITKLLAILGLM	900
DB	841	RLIWLQYFTRABAHQVWVPLNVGRGRDAIILLTCAVHPELIIDITKLLAILGLM	900
QY	901	VLQAGITRVFYFVRAOGLIRACMLVRKVGHHYVQVFMKLGALTGVVYNHLLPLRDWA	960
DB	901	VLQAGITRVFYFVRAOGLIRACMLVRKVGHHYVQVFMKLGALTGVVYNHLLPLRDWA	960
QY	961	HAGRLDVAVEPVVFSAMETKVTIWCADTAACGDIILGLPVSARRGKEIFLGPADSLEG	1020
DB	961	HAGRLDVAVEPVVFSAMETKVTIWCADTAACGDIILGLPVSARRGKEIFLGPADSLEG	1020
QY	1021	QGWELLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVSTATQSLATCTINGVCWT	1080
DB	1021	QGWELLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVSTATQSLATCTINGVCWT	1080
QY	1081	VYHAGSKTLAGPKGPTOMYTNVDLLVHQAPPGARSMTPCSCGSSDLVLTVRHADI	1140
DB	1081	VYHAGSKTLAGPKGPTOMYTNVDLLVHQAPPGARSMTPCSCGSSDLVLTVRHADI	1140
QY	1141	PVRRRGRSGLSPRPVSYLKGSGGGLCPSGHHVGVFRAAVCTRGVAKAVDFIPVES	1200
DB	1141	PVRRRGRSGLSPRPVSYLKGSGGGLCPSGHHVGVFRAAVCTRGVAKAVDFIPVES	1200
QY	1201	METTMRSPVFTDNSTPPPAVPTTFQVAHLHAPTSGSKTKVPAAYAAQGYKVLVLPNSVA	1260

; LENGTH: 2985
; TYPE: PRT
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by nucleotides 2077-11121 of SEQ ID
; OTHER INFORMATION: NO:39
US-10-259-275-40

Query Match 92.5%; Score 14801.5; DB 4; Length 2985;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 2764; Conservative 89; Mismatches 122; Indels 45; Gaps 5;

QY 1 MSTNPKPQRTKXNTNRRPQDVKPCGGQIVGGVYLLPREGPRLGVRATKASERSQPRG 60
Db 1 MSTNPKPQRTKXNTNRRPQDVKPCGGQIVGGVYLLPREGPRLGVRATKASERSQPRG 60

QY 61 RRQIPKARRPEGRAPQGPWPLVYNEGLWAGWLLSPRGRPSMGPTDPRRRSRNLG 120
Db 61 RRQIPKARRPEGRAPQGPWPLVYNEGLWAGWLLSPRGRPSMGPTDPRRRSRNLG 120

QY 121 KVIDTLTCGADLMGYIPLVGA PLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGADLMGYIPLVGA PLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSIIVYEAADVIMHTPGCVPCQEGNSRCWV 240
Db 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSIIVYEAADVIMHTPGCVPCQEGNSRCWV 240

QY 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCFSAMYVDGLCGSIFLYSQLFTSPRRHET 300
Db 241 ALPTTLAARNATPTTIRRHVDLLVGAALCSAMYVDGLCGSVFLVSQLFTSPRRHAT 300

QY 301 VQDCNCSIYPCHVSGHRMAMDMNWSPTTALVVSQQLLRIPQAVDMVAGAHGVLAGLA 360
Db 301 LQDCNCSIYPGHASGHRMAMDMNWSPTTALVVSQQLLRIPQAVDMVAGAHGVLAGLA 360

QY 361 YYSVGNWAKVLIIVALLFAGVDGTHETTRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW 420
Db 361 YYSMAGNWKVLIIVMLLFAGVDGHTLTGGAARLTSGFAGLTTPGPSQRIQLINTNGSW 420

QY 421 HINRTALNCNDSLOTGFPAALFYAHKFNSSCCPERMASCRIPDMFAQGWGPITYTKPNSS 480
Db 421 HINTALNCNDSLOTGFPAALFYAHKFNSSCCPERMASCRIDKFDQGWGPITYVAETKD 480

QY 481 -DQRPYCHVYAPRPGVVPASQVCGPVYCFTPSPVVGTTDRSGVPTYSMEGETDVMLL 539
Db 481 PDQRPYCHVYPPQCGGIVPASQVCGPVYCFTPSPVVGTTDRLGNPTYSMEGETDVMLL 540

QY 540 NNTRPQGNWPGCTWMSNTGTCTCGGPPCNIGGVGNRTLCPTDCPKHPEATYTKGS 599
Db 541 NNTRPQGNWPGCTWMSNTGTCTCGAPPNCNIGGVGNNTLTCTPTDCPKHPEATYSKGS 600

QY 600 GPMWTPRCLVDYVRLWHYPCITNFSIPKVMYVGGVEHRLNAACNTRGRCNLEDRDR 659
Db 601 GPMWTPRCMDYVRLWHYPCITNFSIPKVMYVGGVEHRLNAACNTRGRCNLEDRDR 660

QY 660 SELSPLLISTTEWQILPCAFITLPAISTGLIHLHQNIVDVQYLVGVSFAVSPFAIKWEYI 719
Db 661 SELSPLLISTTEWQVLPSTLPAISTGLIHLHQNIVDVQYLVGVSFAVSPFAIKWEYV 720

QY 720 LLLFLLADARVCACLWMMLLIAQAEAALENLVNNAASVAGAHGILSFLVFFCAAWYIK 779
Db 721 VLLFLLADARVCACLWMMLLIAQAEAALENLVNNAASVAGAHGILSFLVFFCAAWYIK 780

QY 780 GRLLPAGAAFYGVWPLILLILLALPPRAYALDRMAASCGGAVLVGLFTLSPYKVF 839
Db 781 GRLVGAAYAFYGAWPLLLLLLTLPPLAYANDREMAASCGGAVFVGLALLTSPYKVF 840

QY 840 TRLIWLQYFTRAEAHQVWVPLNVRGGRDAILLTCAVHPELIFDITKLLAILGLPL 899
Db 841 ARLLMWLQYLTTRAEAHQVWVPLNVRGGRDAILLTCAVHPELIFDITKLLAILGLPL 900

QY 900 MVLQAGITRVFVRAQGLIRACMLVRKVAGGHYVQVMFKLALGTGTYYVYNHLLTPLRDW 959

Db 901 MVLQAGITRVFVRAQGLIRACMLVRKVAGGHYVQVMFKLALGTGTYYVYNHLLTPLRDW 960
QY 960 AHAGLRDLAVAVEPVVFSAMETKVITWGDATTAACGDIILGLPVSARRGKEIFIGPADSLE 1019
Db 961 AHAGLRDLAVAVEPVVFSAMETKVITWGDATTAACGDIILGLPVSARRGKEIFIGPADSLV 1020
QY 1020 GQGWRLALPITAYSQQTGRGLVCIIITSLTGRDNQVEGEVQVVTATQSFATCINGVCW 1079
Db 1021 -----RDKNQVEGEVQVVTATQSFATCINGVCW 1050
QY 1080 TVYHGAGSKTLAGPKGPIQMTYTNVDLVLVQWQAPFGARSMTPCSCGSSDLYLVTRHADY 1139
Db 1051 TVYHGAGSKTLAGPKGPIQMTYTNVDLVLVQWQAPFGARSLTPTCTCGSSDLYLVTRHADY 1110
QY 1140 IPVRRGDSRGSLLSPRPVSYLKGSSGGLLCPSHGVHGVFRAAVCTRGVAKAVDFIPVE 1199
Db 1111 IPVRRGDSRGSLLSPRPVSYLKGSSGGLLCPSHGVHGVFRAAVCTRGVAKAVDFIPVE 1170
QY 1200 SMETTWRSPVFTDNSTPPRAVPOTFOVAHLHAPTSGSKSTKVPAAVAAQGVKVLVLPNSVA 1259
Db 1171 SMETTWRSPVFTDNSTPPRAVPOTFOVAHLHAPTSGSKSTKVPAAVAAQGVKVLVLPNSVA 1230
QY 1260 ATLGFGAYMSKARGIDPNIRTVGRTITTTGGSITYSTYGFELADGCGSGGAYDIIICDECH 1319
Db 1231 ATLGFGAYMSKARGIDPNIRTVGRTITTTGAPITYSTYGFELADGCGSGGAYDIIICDECH 1290
QY 1320 STDSTTILGIGTVLDOAETAGARLVVLTATPPGCVTVPHNPIEEIGLSNNGEIPFFYGA 1379
Db 1291 STDSTTILGIGTVLDOAETAGARLVVLTATPPGCVTVPHNPIEEIGLSNNGEIPFFYGA 1350
QY 1380 IPTAELKGRSHLIPCHSKKCDLAAKLTGLGHNAYAYRGLDVSVPITPGDVVVATDA 1439
Db 1351 IPTAELKGRSHLIPCHSKKCDLAAKLTGLGHNAYAYRGLDVSVPITPGDVVVATDA 1410
QY 1440 LMTGTGDFDSVDICNTCTVTQTVDFSLDPTFTTETTTVPQDAVSRQRGRTGRGSGIY 1499
Db 1411 LMTGTGDFDSVDICNTCTVTQTVDFSLDPTFTTETTTVPQDAVSRQRGRTGRGSGIY 1470
QY 1500 RFVTPGERPSGMFSSVLCYDAGCAWYELTPTAETSRLRAYLNTPLGVPQDHLBFWE 1559
Db 1471 RFVTPGERPSGMFSSVLCYDAGCAWYELTPTAETSRLRAYLNTPLGVPQDHLBFWE 1530
QY 1560 SVFTGLTHIDAHFLSQTQKAGDNFPLYVAYQATVCARQAQPPSPSDQWMLKRLKPTLH 1619
Db 1531 SVFTGLTHIDAHFLSQTQKAGDNFPLYVAYQATVCARQAQPPSPSDQWMLKRLKPTLH 1590
QY 1620 GPTPLLRGLGAVONEVILTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTIGS 1679
Db 1591 GPTPLLRGLGAVONEVILTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTIGS 1650
QY 1680 VVIVGRILISGKPAVVPDREVLVYQEFDEMEECASQLPYIEQGMQALAEQFKQKALGLQTA 1739
Db 1651 VVIVGRILISGKPAVVPDREVLVYQEFDEMEECASQLPYIEQGMQALAEQFKQKALGLQTA 1710
QY 1740 TKQAEAAAPVVEKSWRALETFWAKHKNWFIISGIIYLAGLSTLPCNPAIASIMATASITS 1799
Db 1711 TKQAEAAAPVVEKSWRALETFWAKHKNWFIISGIIYLAGLSTLPCNPAIASIMATASITS 1770
QY 1800 PLTTQNTLLFNILGWWAAQLAPPASAASFVAGIAGAAVGSIGLGVLDILAGYAGV 1859
Db 1771 PLTTQNTLLFNILGWWAAQLAPPASAASFVAGIAGAAVGSIGLGVLDILAGYAGV 1830
QY 1860 AGALVAFKWSGVEPSTEDLVNLLPAILSPCALVGVVCAAILRRHVPGEVQVMNRL 1919
Db 1831 AGALVAFKWSGVEPSTEDLVNLLPAILSPCALVGVVCAAILRRHVPGEVQVMNRL 1890
QY 1920 IAFASRGNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHQNINEDCSTPCSGSWLRD 1979
Db 1891 IAFASRGNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHQNINEDCSTPCSGSWLRD 1950
QY 1980 VMDWICVLTDFKWLQSLPLRPVGPVPLSCORGYGVWFGDGIQMTQTCGCAQIAGHV 2039

Db 1951 VDMWCTVLSDFKTWLQSKLRLPGVFPFLSCQRGYKGVWRGDIHMTTCPGAQIAGHV 2010
QY 2040 KNGSMRIYGPRTCSNTWGTTFPINAYTTGCPSPAPNYSRALWRVAABEYVEVTRVGDF 2099
Db 2011 KNGSMRIIGPKTCSNTWGTTFPINAYTTGCPSPAPNYSKALWRVAABEYVEVTRVGDF 2070
QY 2100 HYVTGWTIDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLREDVTFQVGLNQYLVSQ 2159
Db 2071 HYVTGITTIDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLREDVTFQVGLNQYLVSQ 2130
QY 2160 L-----PCEPEDVTVLTSMLTDPSHITAEAKRLARGSPSSLASSASQLSAPLKAT 2214
Db 2131 LHYRAPVKCP-----LLRDEVVQVGLAQYLVSQSLARGSPSSLASSASQLSAPLSLAT 2185
QY 2215 CTTTH-----DSPDADLIEANLWROEMGNITRVESENKVVILDSFEPHABGDEREISV 2270
Db 2186 CTTTHSYNLDSPVDLIAANLWROEMGNITRVESENKVVVLDSEFPLRABGDENEISI 2245
QY 2271 AAEILKRKRKPPSALPIWARPDYNNPPLLESWKDPDYPVPPVHGCPLPPTKAPPIPPRRK 2330
Db 2246 AAEILKRKRKPPAIPWARPDYNNPPLLESWKNDYVPPVHGCPLPVKAPPIPPRRK 2305
QY 2331 RTVVLTESNVSSALAEALATKTFGSSGSSAVIDSGTATAPDLASDDGDKGSDVESYSSMP 2390
Db 2306 RTVVLTSTSVSSVLAELATKTFGSSSELSAADSQTATAPPDQTSNDGKSDAESCSMP 2365
QY 2391 LEGEPDGLDSGNSVSEEBASEDVCCSMSTWTGALITPCAAEESKLPINLNSILL 2450
Db 2366 LEGEPDGLDSGNSVSEEBASESVCCSMSTWTGALITPCAAEESKLPINALNSULL 2425
QY 2451 RHHNVYATTSASLRQKKVTFDRQLQVLDHYRDLVKEMKAKASTVKAKLSIEBEACKL 2510
Db 2426 RHHNVYATTSASLRQKKVTFDRQLQVLDHYRDLVKEMKAKASTVKAKLSIEBEACRL 2485
QY 2511 TPHPGAKFGYGAQDVNRNLSRAVNHIRSVWEDLLEDTEPIDTIMAKSEVFCVQPEK 2570
Db 2486 TPHPGAKFGYGAQDVNRNLSRAVNHIRSVWEDLLEDTEPIDTIMAKSEVFCVQPEK 2545
QY 2571 GGRKPARLIVFDLGVRCERKALVDVSTLPOAVMGSSYGQYQPKORVEFLVNTWSK 2630
Db 2546 GGRKPARLIVFDLGVRCERKALVDVSTLPOAVMGSSYGQYQPKORVEFLVKAWSK 2605
QY 2631 KCPMGFSYDTRCFDSTVESDIRVEESYQCCDLAPPEARQAIRSLTERLYIGGLTNSKG 2690
Db 2606 KNPMGFSYDTRCFDSTVENDIRVEESYQCCDLAPPEARQAIRSLTERLYIGGLTNSKG 2665
QY 2691 QNCGYRRCASGLVTTSCGNLTLCYLKATAACRAAKLQDCTMLVNGDDLWICESAGTQE 2750
Db 2666 QNCGYRRCASGLVTTSCGNLTLCYLKASAAKRAAKLQDCTMLVNGDDLWICESAGTQE 2725
QY 2751 DAAALRAEATRYASAPGDPQPEYDLELITSCSNVSVAHADSGKRVVYLTRDPTTP 2810
Db 2726 DAAALRAEATRYASAPGDPQPEYDLELITSCSNVSVAHADSGKRVVYLTRDPTTP 2785
QY 2811 LARAAWETARHTPINSGLNIIWYAPTILWARMILMTHFPIISLLAQEKLKALDCCIYGAC 2870
Db 2786 LARAAWETARHTPVNSWGLNIIWYAPTILWARMILMTHFPIISLLAQEKLKALDCCIYGAC 2845
QY 2871 YSIEPLDLPQIIEHLHGLSAFTLHYSYSGEINRVASCLKGLVPLTRWRHRAVSRAKL 2930
Db 2846 YSIEPLDLPQIIEHLHGLSAFTLHYSYSGEINRVASCLKGLVPLTRWRHRAVSRAKL 2905
QY 2931 LSGGGAATCGYLFNWAATKLTPIPAASQLDLSGFVAGYSGGDIYHLSRARPRW 2990
Db 2906 LSGGGAATCGYLFNWAATKLTPIPAASQLDLSGFVAGYSGGDIYHLSRARPRW 2965
QY 2991 FPLCLLLLSVGVGIYLLPNR 3010
Db 2966 FPLCLLLLSVGVGIYLLPNR 2985

RESULT 4

, US-11-006-313-40

; Sequence 40, Application US/11006313
; Publication No. US20050153281A1
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley M.
; APPLICANT: Yi, Minkyung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0121
; CURRENT APPLICATION NUMBER: US/11/006,313
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 10/259,275
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 09/747,419
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/325,236
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 2985
; TYPE: PRT
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by nucleotides 2077-11121 of
; OTHER INFORMATION: SEQ ID NO:39
; US-11-006-313-40

Query Match 92.5%; Score 14801.5; DB 6; Length 2985;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 2764; Conservative 89; Mismatches 122; Indels 45; Gaps 5;

QY 1 MSTNPKPQRTKRNTRRRPQDVKFPFGGQIYGVGVLLPRRGPRLGVATRAKASRSQRG 60
Db 1 MSTNPKPQRTKRNTRRRPQDVKFPFGGQIYGVGVLLPRRGPRLGVATRAKTSERSQRG 60
QY 61 RRQPTPKARRPEGRAWAQPYPWPLYGNEGEGWAGWLLS PRGSRPSWGTDPDRRSRNLG 120
Db 61 RRQPTPKARRPEGRAWAQPYPWPLYGNEGEGWAGWLLS PRGSRPSWGTDPDRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAFLGGAARALAHGVRLVEDGVNYATGNLPGCSFIFLLA 180
Db 121 KVIDTLTCGLADLMGYIPLVGGPLGGAARALAHGVRLVEDGVNYATGNLPGCSFIFLLA 180
QY 181 LLSCLTIPASAYEVNRVNSGIYHVNTDCSNSISIVYEAADVIMHTPCGVCPCVQEGNSSRCWV 240
Db 181 LLSCLTVPASAEVRNASGVYHVNTDCSNSISIVFEADLIMHTPCGVCPCVREGNSSRCWV 240
QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYGLDLCGSIFLVSQLFTFSRRHET 300
Db 241 ALTPTLAARNATIPPTTIRRHVDLLVGAALCSAMVYGLDLCGSVFLVSQLFTFSRRHAT 300
QY 301 VQDCNCSYIPGHVSGHRMADMMNNSPTTALVVSQLLRIPOAVVDMVAGAHGWVLGALA 360
Db 301 LQDCNCSYIPGHASGHRMADMMNNSPTTALVVSQLLRIPOAVIDMVAGAHGWVLGALA 360
QY 361 YYSVMGVNKAIVLIVALLFAGVDGTHTTGRVAGHTTSFTSLFSSGASQKIQLVNTNGSW 420
Db 361 YYSMAGNKAIVLIVMLFAGVDGHTLTGTGHAARLTSGFAGLFTPGPSQRIQLINTNGSW 420
QY 421 HINRTALNCNDSLOTFEALFYAHKFNSSCGPERMASCRPIDWFAQCGWGPITYTKPNSS 480
Db 421 HINRTALNCNDSLOTFGLAALFYAHRFNSSCGPERMASCRSIDKFDQWGPITYAETPKD 480
QY 481 -DQRPYCMHYAPRCGVVPASQVCGPVYCFPTSPVVVGTTRDRSGVPTYSWGENETDVMLL 539
Db 481 PDQRPYCMHYPPQCGIIVPAQVCGPVYCFPTSPVVVGTTRDELGNPTYSWGENETDVMLL 540
QY 540 NNTRPQGNWFGCTWMNSTGTFTKTCGPPCNIGGVGNRTLICPTDCFRKHPEATYTKGSS 599
Db 541 NNTRPQGNWFGCTWMNSTGTFTKTCGAPPNCNIGGVGNNTLTCPTDCFRKHPEATYTKGSS 600

600 QY GPMLTRCLVDYRULWHYPTCLNFSIFKVRMTVGGVEHRLNAACNWTGRGERCNLEDDR 659
601 Db GPMLTRPCMVDPYRULWHYPTCNFSIFKVRMTVGGVEHRLNAACNWTGRGERCNLEDDR 660
660 QY SELSPULLSTTEWQILPCAFRTLLPALSTGLIHLHONIVDOXYLYGVGSAPVSAIKWEYI 719
661 Db SELSPULLSTTEWQILPCAFRTLLPALSTGLIHLHONIVDOXYLYGVGSAPVSAIKWEYV 720
720 QY LILFLLLADARVCACILMMMLLIAQAEALLENVLNAASVAGAHGILSLFVFFCAAWYIK 779
721 Db LILFLLLADARVCACILMMMLLIAQAEALLENVLNAASVAGAHGILSLFVFFCAAWYIK 780
780 QY GRILAPGAAYAFYGWPLLLLLALPPRAYALDREMAASCGGAYLVGLVFLTSLPYKVF 839
781 Db GRILVPGAAYAFYGAWPLLLLLLTLPPrAYAMDRMAASCGGAVFVGLALLTSLPYKVF 840
840 QY TRLIWMLQYPIITRAEAMHOMVWPLNVRGORDAIILITCAVHPELIFDITKLLAILG 899
841 Db ARLLWMLQYLIITRAEALHVMWPLNVRGORDAIILITCAVHPELIFDITKLLAILG 900
900 QY MVLQAGITRVPYFVRAOGLIRACMLVRKVGAGHYVQMAFMRGALTGTYYVYNHLTPLRDW 959
901 Db MVLQAGITRVPYFVRAOGLIRACMLVRKVGAGHYVQMAFMRGALTGTYYVYNHLTPLRDW 960
960 QY AHAGLRDLAVAVPVPVFSAMETKVIITWADTAACGDIILGLPVSARRGKFIIFLGPADSLE 1019
961 Db AHAGLRDLAVAVPVPVFSAMETKVIITWADTAACGDIILGLPVSARRGKFIIFLGPADSLE 1020
1020 QY GOGWRLAPITAYSQTRGVLCIITSLTGRDNQVEGEVQVVSATQSFATCINGVCW 1079
1021 Db GOGWRLAPITAYSQTRGVLCIITSLTGRDNQVEGEVQVVSATQSFATCINGVCW 1080
1080 QY TVYHAGSKTLAGKGPITOMYTNVDLDLVGWAQPPGARSMTPCSCSSDLYLVRHADV 1139
1081 Db TVYHAGSKTLAGKGPITOMYTNVDLDLVGWAQPPGARSMTPCSCSSDLYLVRHADV 1140
1140 QY IPVRRCDRSGSLLSPVPSVYKSGSGPLLCPSGHVGVFRAVCTRGVAKAVDRTIPE 1199
1141 Db IPVRRCDRSGSLLSPVPSVYKSGSGPLLCPSGHVGVFRAVCTRGVAKAVDRTIPE 1170
1200 QY SMETTMRSPTFTNSTPAPVQFQVAHLHAPTGSKGSTKVPAAAYAAQGYKVLVNP 1259
1171 Db SMETTMRSPTFTNSTPAPVQFQVAHLHAPTGSKGSTKVPAAAYAAQGYKVLVNP 1230
1260 QY ATLGFGAYMSKAHGDINIRITVTGGSITYSTYTKFLADGCGSGGAYDIICDECH 1319
1231 Db ATLGFGAYMSKAHGDINIRITVTGGSITYSTYTKFLADGCGSGGAYDIICDECH 1290
1320 QY STDSSTTLIGTGLVLDQAEATAGARLVLATATPPGVSVPVPHNIBEIGLSNNGEIPFYGKA 1379
1291 Db STDSSTTLIGTGLVLDQAEATAGARLVLATATPPGVSVPVPHNIBEIGLSNNGEIPFYGKA 1350
1380 QY IPTBAIKGRHLPFCHSKKKDELAALKTGLGLNAVAYRGLDVSVPPIPGDVVVATDA 1439
1351 Db IPTBAIKGRHLPFCHSKKKDELAALKTGLGLNAVAYRGLDVSVPPIPGDVVVATDA 1410
1440 QY LMTGTGDFSDVIDCNTVQTVDFSLDPTFTIETTVPODAVSRORRGTRGRSGIY 1499
1411 Db LMTGTGDFSDVIDCNTVQTVDFSLDPTFTIETTVPODAVSRORRGTRGRSGIY 1470
1500 QY RFVTPGRSPSGMFSSVLCYDAGCAWYELTPAETSURLRAYLNTPLPVCQDHLFEWE 1559
1471 Db RFVTPGRSPSGMFSSVLCYDAGCAWYELTPAETSURLRAYLNTPLPVCQDHLFEWE 1530
1560 QY SVFTGLTHIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSDQWQKCLIRLKP 1619
1531 Db SVFTGLTHIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSDQWQKCLIRLKP 1590
1620 QY GPTPLLRGLAVQNEVILTHPITKYINACMSADLEVTVSTWVNGVLAALAAAYCLTGS 1679
1591 Db GPTPLLRGLAVQNEVILTHPITKYINACMSADLEVTVSTWVNGVLAALAAAYCLTGS 1650

1680 QY VVIVGRIILISGKPAVVPDREVLVYQREDEMEECASOLPYIEQGMQLAEQFKOKALGLQTA 1739
1651 Db VVIVGRIILISGKPAVVPDREVLVYQREDEMEECASOLPYIEQGMQLAEQFKOKALGLQTA 1710
1740 QY TKQAAAAAIVBESKWRALFTFWAKHWNFIISGTQYLAGLSTLPGNPAIASLMAFTASITS 1799
1711 Db TKQAAAAAIVBESKWRALFTFWAKHWNFIISGTQYLAGLSTLPGNPAIASLMAFTASITS 1770
1800 QY PLTTQNTLLFNILGWWAAQLAPPSAASAFVAGIAGAAVGSIGLQKLVLDIILAGVAGV 1859
1771 Db PLTTQNTLLFNILGWWAAQLAPPSAASAFVAGIAGAAVGSIGLQKLVLDIILAGVAGV 1830
1860 QY AGALVAFKMSGEVPESTEDLVNLLPAILSPGALVGVCAAILRRHVGEGBGVQWNNRL 1919
1831 Db AGALVAFKMSGEVPESTEDLVNLLPAILSPGALVGVCAAILRRHVGEGBGVQWNNRL 1890
1920 QY IAFASRGNHVSPTHVVPESDAARVTOILSSLTITQLLKRLHOWINEDCSTPCSGSWLRD 1979
1891 Db IAFASRGNHVSPTHVVPESDAARVTOILSSLTITQLLKRLHOWINEDCSTPCSGSWLRD 1950
1980 QY VMDWICTVLTDFTKWLQSKLLPRLPGVPFLSCQGYKGVWRGDMQTTCTPCGAQIAGHV 2039
1951 Db VMDWICTVLTDFTKWLQSKLLPRLPGVPFLSCQGYKGVWRGDMQTTCTPCGAQIAGHV 2010
2040 QY KNGSMRIVGRTCSNTHWGTFFINAYTTGCTPSPAPNYSRALMRVAABEYEVTRVGDF 2099
2011 Db KNGSMRIVGRTCSNTHWGTFFINAYTTGCTPSPAPNYSRALMRVAABEYEVTRVGDF 2070
2100 QY HYVTGTTDNVCKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEDVTFQVGLNOYLVSQ 2159
2071 Db HYVTGTTDNVCKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEDVTFQVGLNOYLVSQ 2130
2160 QY L-----PCPEPDDVTVLTSMLTDPSHITAEAKRRLARGSPPLSSASSASOLSAPSLKAT 2214
2131 Db L-----PCPEPDDVTVLTSMLTDPSHITAEAKRRLARGSPPLSSASSASOLSAPSLKAT 2185
2215 QY CTTTHH-----DSPADLIEANLLWRQEMGNIITRVESENKVIILDSFEPLHAEGDERISV 2270
2186 Db CTTTHH-----DSPADLIEANLLWRQEMGNIITRVESENKVIILDSFEPLHAEGDERISV 2245
2271 QY ABEILKRSKFPSPALFIWARPDYNNPPLLESWKDPDYVPPVHVHGCPLPPTKAPDIPPRRK 2330
2246 Db ABEILKRSKFPSPALFIWARPDYNNPPLLESWKDPDYVPPVHVHGCPLPPTKAPDIPPRRK 2305
2331 QY RTVVLTESNVSSLAELATKTFGSSSSASVDSGTATALPDASDDGDKGSDVESYSWMP 2390
2306 Db RTVVLTESNVSSLAELATKTFGSSSSASVDSGTATALPDASDDGDKGSDVESYSWMP 2365
2391 QY LECEPDDPLSDGSSWSTVSEASEDVVCCSMSVTWTGALITPCAABESKLPINPLNSLL 2450
2366 Db LECEPDDPLSDGSSWSTVSEASEDVVCCSMSVTWTGALITPCAABESKLPINPLNSLL 2425
2451 QY RHNHMYATTSRASLRQKVTDFRLQVLDHDDYRDVLKEMKAKASTVKAKLLSIEEACKL 2510
2426 Db RHNHMYATTSRASLRQKVTDFRLQVLDHDDYRDVLKEMKAKASTVKAKLLSIEEACKL 2485
2511 QY TTPHSAKSGYGAQVNRNLSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVFCVQPEK 2570
2486 Db TTPHSAKSGYGAQVNRNLSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVFCVQPEK 2545
2571 QY GGRKPARLIIVPDLGVVRCEKMALYDVVSTLPOAVMGSSYGFQYSKQRFVFLVNTWKS 2630
2546 Db GGRKPARLIIVPDLGVVRCEKMALYDVVSTLPOAVMGSSYGFQYSKQRFVFLVNTWKS 2605
2631 QY KCPMGSYDTRCFDSTVTESDIRVESIYOCDDLAPEARQAIIRSLTERLYIGGPLTNSKG 2690
2606 Db KCPMGSYDTRCFDSTVTESDIRVESIYOCDDLAPEARQAIIRSLTERLYIGGPLTNSKG 2665
2691 QY QNCGYRCRASGVLTTCYKATATAACRAAKLQDCTMLVNGDDLVVICESAGTQE 2750
2666 Db QNCGYRCRASGVLTTCYKATATAACRAAKLQDCTMLVNGDDLVVICESAGTQE 2725
2751 QY DAAALRAFTAMTRYAPPDPPQPEYDLEBLITSCSSNVSAHDASGRVYVYLTRDPTTP 2810


```
Db 2726 DAASLRVETAMTRYSAPEGDLUPQEPYDLLETSCSNVSAHDASGRVYVYLRDPTTP 2785
Qy 2811 LARAAMETARHTPIINSWGLNIIMYAPTILWARMILMTHFFSILLAOEQLKALDCCIYGAC 2870
Db 2786 LARAAMETARHTPIINSWGLNIIMYAPTILWARMILMTHFFSILLAOEQLKALDCCIYGAC 2845
Qy 2871 YSIEPLDLPQIIERLHGLSAFTLHYSFGEINRVASCLRLKGLGVPPLRTWRHRAERSVRAL 2930
Db 2846 YSIEPLDLPQIIERLHGLSAFSLHSYSFGEINRVASCLRLKGLGVPPLRVWRHRAERDVRAL 2905
Qy 2931 LSQGRATCGRYLFWNAVRTKLTPIPAASQDLSCGFVAGYSGGDIYHSLSPARPRW 2990
Db 2906 LSQGRATCGRYLFWNAVRTKLTPIPAASQDLSCGFVAGYSGGDIYHSLSPARPRW 2965
Qy 2991 FPLCLLLLSVGVIYLLPNR 3010
Db 2966 FPLCLLLLSVGVIYLLPNR 2985

RESULT 5
US-09-742-659-4
; Sequence 4, Application US/09742659
; Patent No. US20010034019A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallio, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3011
; TYPE: PRF
; ORGANISM: Hepatitis C virus
US-09-742-659-4

Query Match      88.5%; Score 14162.5; DB 3; Length 3011;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 2579; Conservative 217; Mismatches 214; Indels 1; Gaps 1;

Qy 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVYLLPRGRLGVRATRKASERSQPRG 60
Db 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVYLLPRGRLGVRATRKTSERSQPRG 60
Qy 61 RQPIPKARRPEGRAWAQCYGWPPLYGNEGGLWAGWLLSPGRSPSGWPTDPRRSRNUG 120
Db 61 RQPIPKARRPEGRWAQCYGWPPLYGNEGCGWAGWLLSPGRSPSGWPTDPRRSRNUG 120
Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTIPASAEVRNVSGIYHVNTDCNSSIYVEAADVIMHTPGCVPCQEGNSRCRW 240
Db 181 LLSCLTIPASAEVRNVSGIYHVNTDCNSSIYVEAADVIMHTPGCVPCQEGNSRCRW 240
Qy 241 ALTPTLAARNASVPTTIRRHVDLVAGTAFCSAMVVDGLCGSIFLVSQLFTFSRRHET 300
Db 241 AVTPTVATRDGKLTQLRRHDLVLSATLCSALYVGDLCGSFVLVGQLFTFSRRHET 300
Qy 301 VDCNCSIPYGHVSGHRMAWMMNWSPTTALVVSQLLRIPQAVVDMVAGAHGWVLGIA 360
Db 301 TQDCNCSIPYGHVSGHRMAWMMNWSPTTALVVSQLLRIPQAVVDMVAGAHGWVLGIA 360
```

```
Qy 361 YYSVMGNWAKVLIIVALLPAGVDGETHHTGRVAGHTTSGFTSLFSSGASQKLOLVNTNGSW 420
Db YYSVMGNWAKVLIIVALLPAGVDGETHHTGRVAGHTTSGFTSLFSSGASQKLOLVNTNGSW 420
Qy 421 HINETALNCNDSLOTGFPAALFYAHKENSSCPERMASCRPIDWFAQGWGPITVTKPNSS 480
Db HINETALNCNDSLOTGFPAALFYAHKENSSCPERMASCRPIDWFAQGWGPITVTKPNSS 480
Qy 481 DORYCHWYAPRCPGVPVSPASOVCGPVVCFPTSPVVVGTDRSGVPTYSWGENEDVMLIN 540
Db DORYCHWYAPRCPGVPVSPASOVCGPVVCFPTSPVVVGTDRSGVPTYSWGENEDVMLIN 540
Qy 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVNRTLICPTDCFRKHPKPYATYKCSG 600
Db NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVNRTLICPTDCFRKHPKPYATYKCSG 600
Qy 601 PWLTFRCLVDYPRYLWHYPCTLNFISIFKVRMYGVGVEHRLNAACNWRGERCNLEDRDRS 660
Db PWLTFRCLVDYPRYLWHYPCTLNFISIFKVRMYGVGVEHRLNAACNWRGERCNLEDRDRS 660
Qy 661 ELSPLLLSTTEWOJILPCAFTTLPALSTGLIHLHONIVDVQVLYGVGSFAVFAIKWEYIL 720
Db ELSPLLLSTTEWOJILPCAFTTLPALSTGLIHLHONIVDVQVLYGVGSFAVFAIKWEYIL 720
Qy 721 LLFLLADARVCACILWMLLIQAEEALENIJVLNAASVAGAHGILSLFVFFCAAWYIKG 780
Db LLFLLADARVCACILWMLLIQAEEALENIJVLNAASVAGAHGILSLFVFFCAAWYIKG 780
Qy 781 RLAPGAAYAFYGVNPLLLLLLALPPRAYALDREMAASCGGAVLVGLVFLTSPYKYKFLT 840
Db RLAPGAAYAFYGVNPLLLLLLALPPRAYALDREMAASCGGAVLVGLVFLTSPYKYKFLT 840
Qy 841 RLWVLOVYFTRAEAHMQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGLPM 900
Db RLWVLOVYFTRAEAHMQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGLPM 900
Qy 901 VLOAGITRVYPVRAQGLIRACMLVRKVAGGHYVQVPMFKLGALTGTYYVYVHLTPLRDA 960
Db VLOAGITRVYPVRAQGLIRACMLVRKVAGGHYVQVPMFKLGALTGTYYVYVHLTPLRDA 960
Qy 961 HAGLRDLAVAVEPVVFSAMETKVITWGDATACGDIILGLPVSARRGKEIFLPGADSLG 1020
Db HAGLRDLAVAVEPVVFSAMETKVITWGDATACGDIILGLPVSARRGKEIFLPGADSLG 1020
Qy 1021 QGWRLLAPITAYSOOTRGVLCIITSLTRDKNOVEGEVQVSTATQSFATCINGVCWT 1080
Db QGWRLLAPITAYSOOTRGVLCIITSLTRDKNOVEGEVQVSTATQSFATCINGVCWT 1080
Qy 1081 VYHAGSKTLAGPKPITQMTYNTVDLDLVGWOAPPGARSMTPCSCSSDLVLTVRHADVI 1140
Db VYHAGSKTLAGPKPITQMTYNTVDLDLVGWOAPPGARSMTPCSCSSDLVLTVRHADVI 1140
Qy 1141 PVRRGDSRGLSLLSPRVSYLYKSGSGGFLPCSPGHVGVFRAAVCTRGVAKAVDFIPVES 1200
Db PVRRGDSRGLSLLSPRVSYLYKSGSGGFLPCSPGHVGVFRAAVCTRGVAKAVDFIPVES 1200
Qy 1201 METTWRSPVFTDNSTPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGVKVLVNPVAA 1260
Db METTWRSPVFTDNSTPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGVKVLVNPVAA 1260
Qy 1261 TLGFCAYMSKAGHDIPNIRTGVRTITTCGSTYTSYTGKFLADGGCGGAYDIIICDECHS 1320
Db TLGFCAYMSKAGHDIPNIRTGVRTITTCGSTYTSYTGKFLADGGCGGAYDIIICDECHS 1320
Qy 1321 TDSTTILGTVLDQAEATAGARLVVLTATATPPGSVTVPHPNIEEIGLSNNGEIPFYGKAI 1380
Db TDSTTILGTVLDQAEATAGARLVVLTATATPPGSVTVPHPNIEEIGLSNNGEIPFYGKAI 1380
Qy 1381 PIEATIKGRHLLIFCHSKKKKDELAALKLGLGLNVAAYYRGLDVSVIPIPGDWWWATDAL 1440
Db PIEATIKGRHLLIFCHSKKKKDELAALKLGLGLNVAAYYRGLDVSVIPIPGDWWWATDAL 1440
```



```
QY 1441 MTGFTGDFSDVIDCNTCVTOTVDFSLDPTTTIETTTPQDAVRSQRGRGTGRGIYR 1500
Db 1441 MTGFTGDFSDVIDCNTCVTOTVDFSLDPTTTIETTTPQDAVRSQRGRGTGRGIYR 1500
QY 1501 FVTPGERPSGWFSSVLCCEYDAGCAWELTPAETSURLRAYLNTPLCPVQCDHLEWES 1560
Db 1501 FVAPGERPSGWFSSVLCCEYDAGCAWELTPAETIVRLRAYNTPLCPVQCDHLEWEG 1560
QY 1561 VFTGLTHIDAHFLSQTQAGDNFPYLVAQYATCARAQAPPPSDQWQKCLIRLKP TLHG 1620
Db 1561 VFTGLTHIDAHFLSQTQSGENFPYLVAQYATCARAQAPPPSDQWQKCLIRLKP TLHG 1620
QY 1621 PTLPLRLGAVQNEVILTHITIKYIMACMSADLEVTSTWLVGGVLAALAAALCLITGTV 1680
Db 1621 PTLPLRLGAVQNEVILTHITIKYIMTMSADLEVTSTWLVGGVLAALAAALCLSTGCV 1680
QY 1681 VIVGRITLSKPAVVPDREVLVQEFDEMERCASOLPYIEOGMOLAEQFKALGLLQAT 1740
Db 1681 VIVGRIVLSGKPAIIPREVLVQEFDEMEBCSOHLPIYIEOGMLAEQFKALGLLQAT 1740
QY 1741 KQAEAAAPVYESKRALJETFWAKHMNFISGIOYLAGLSTLPGNPATIASLMAFTASITSP 1800
Db 1741 RQAEVITPAVQTNWQKLEVFQAKHMNFISGIOYLAGLSTLPGNPATIASLMAFTAVTSP 1800
QY 1801 LTTQNTLLFNILGWAAQALAPPASAFVAGAGIAGAAVSGIGLKVLDILAGYGAGVA 1860
Db 1801 LTTGQTLLEFNILGWAAQALAAAPGAATAFVAGAGLAGAAIGSVGLKVLVDILAGYGAGVA 1860
QY 1861 GALVAFKMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEQAVQWNNRLI 1920
Db 1861 GALVAFKMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEQAVQWNNRLI 1920
QY 1921 AFASRGNHVSPTHVVPSPDAAARVTOILSSITITQLLKRHLHOMINEDCSTPCSGSLRDV 1980
Db 1921 AFASRGNHVSPTHVVPSPDAAARVAILSSITVTOILLRRLHOMISSECTPCSGSLURDI 1980
QY 1981 WDMTCTVLTDFKTLQSKLPLRPGVFPFLSCQRYGKGVWRGDGIMQITCPCGAQIAGHVK 2040
Db 1981 WDMICEVLSDFKTLWAKLNPOLPGIFVSCQRYGKGVWRGDGIMHTRCHCGAEITGHVK 2040
QY 2041 NGSMRIVGPRCTSNWHTGTFPINAATTGCTPPAPNYSRALMRVAAEEVVEYTRVGDHF 2100
Db 2041 NGTMRIVGPRCTSNWSTGTFPINAATTGCTPLPAPNYPKALMRVSAEEVVEIRRVGDHF 2100
QY 2101 YVTQMTDNKCPQVAPAEFFTEVDGURLHRYAPACKPLLRDVTFOVGLNQVLGSOL 2160
Db 2101 YVSGMTTDLNKCPCQIPSPPEFTLDGURLHRYAPPCKPLLRREVSFRVLGHEYPVGSOL 2160
QY 2161 PCEPEPDVTLTSMLTDPSSHITAEAKRLARGSPPSILASSASQLSAPSLKATCTTHHD 2220
Db 2161 PCEPEPDVAVLTSMLTDPSSHITAEAGRLARGSPPSWASSASQLSAPSLKATCTANHD 2220
QY 2221 SPDAELIEANLLRWQMGNTIRVESENKVILDSFELPAEGDEREISVAAELIRKSRK 2280
Db 2221 SPDAELIEANLLRWQMGNTIRVESENKVILDSFELPAEEDEREVSYPAEILIRKSR 2280
QY 2281 FPSALPIWARDYNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTSNV 2340
Db 2281 FAPALPVWARDYNPPLVETWKDPDYBPPVHVHGCPLPPPPRPPPPVPPRKKRTVVLTSNL 2340
QY 2341 SSALAEALATKTFSGSGSANDSGTATALPDLASDGDGKSGSVESYSMPPLLEGFPDPL 2400
Db 2341 STALAEALATKTFSGSSSTSGITGDNNTTSSPEAPSGCPDSDSVESYSMPPLLEGFPDPL 2400
QY 2401 SDGWSVTSEEA-SEDVVCCSWSVTWTCALITPCAABESKLPIPNLSLIRHNNVYAT 2459
Db 2401 SDGWSVTSSGADTEDVVCCSWSVTWTCALITPCAABEQKLPINALSLSLIRHNNLYST 2460
QY 2460 TSSASLRQKQKVTDFRLQVLDHVDVLEKEMAKASTVAKLLSIEACKLTPPHSAKSK 2519
Db 2461 TSSACQKQKVTDFRLQVLDHVDVLEKEMAKASTVAKLLSIEACKLTPPHSAKSK 2520
QY 2520 FGYGAKDVNRILSRVNHIRSWEDLLDLETPTIDTTIMAKSEVFCVQPEKGRKPARLI 2579
```

```
Db 2521 FGYGAKDVNRCHARKAVAHINSVWMDLLEDSVTPIDTTIMAKNEVFCVQPEKGRKPARLI 2580
QY 2580 VFPDLGVRVCEKXKALVDVYSTLPQAVMGSSYGFQYSPKORVEFLVNTWKSKKCPMGFSYD 2639
Db 2581 VFPDLGVRVCEKXKALVDVYSTLPQAVMGSSYGFQYSPGQORVEFLVQAMSKSKTTPMGFSYD 2640
QY 2640 TRCFDSTVTESDIRVESIYQCCDLAPAPARQAIRSLTERLYIGGPLTNSGQNGCYRRCR 2699
Db 2641 TRCFDSTVTESDIRTEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCR 2700
QY 2700 ASGLVLTSCGNTLTJCYLKATAACRAAKLODCTMLVNGDDLVICESAGTQEDAAALRAFT 2759
Db 2701 ASGLVLTSCGNTLTJCYKARAAACRAAGLODCTMLVCGDDLVICESAGVQEDAAALRAFT 2760
QY 2760 EAMTRISAPPGDPPOPEYDLELITSCSSNVSAHDASGKRVYVLTDRPTTPLARAAWETA 2819
Db 2761 EAMTRISAPPGDPPOPEYDLELITSCSSNVSAHDGAGKRVYVLTDRPTTPLARAAWETA 2820
QY 2820 RHTPINSWLGNIIMYAPTILWARMILMTHFFSILLAQEQLEKALDCQIYGACYIEBPLDLP 2879
Db 2821 RHTPVNSWLGNIIMFAPTILWARMILMTHFFSVLIARDQLEQALNCEIYGACYIEBPLDLP 2880
QY 2880 QIITERLHGLSAFTLHYSYSGEINRVASCLKGLVPLRTWRHRARSVRAKLLSOGGRAAT 2939
Db 2881 PIITORLHGLSAFSLHYSYSGEINRVAAACLKGLVPLRLAWHRARSVRALLSRGGRAAI 2940
QY 2940 CGRYLFENWAVRTKLKLTPIPAASOLDLSGWFVAGYSGDIYHLSLRARPRWFPCLCLLLS 2999
Db 2941 CGRYLFENWAVRTKLKLTPIPAAGRLDLSGMFTAGYSGDIYHLSVSHARPRWFCFCLLLA 3000
QY 3000 VGVGIYLLPNR 3010
Db 3001 AGVGIYLLPNR 3011

RESULT 6
US-09-891-894-3
; Sequence 3, Application US/09891894
; Publication No. US20030013081A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Paul
; APPLICANT: Maddon, Paul
; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS INF
; FILE REFERENCE: 2048/64896/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/891,894
; CURRENT FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: hepatitis c virus
US-09-891-894-3

Query Match      88.3%; Score 14143.5; DB 3; Length 3011;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 2576; Conservative 218; Mismatches 216; Indels 1; Gaps 1;

QY 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRAATRKASERSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRAATRKTSERSQPRG 60
QY 61 RROPICKARPEGRGAWAQPGYPMPLVYCNELGWAGLLSPRSGRPSWGTDPDRRRSRNLG 120
Db 61 RROPICKARPEGRGAWAQPGYPMPLVYCNELGWAGLLSPRSGRPSWGTDPDRRRSRNLG 120
QY 121 KVITDLTCGFADLMGYIPLVGA PLGGAARALAHGVRVLEQVNYATGNLPCGSFSIFLLA 180
Db 121 KVITDLTCGFADLMGYIPLVGA PLGGAARALAHGVRVLEQVNYATGNLPCGSFSIFLLA 180
QY 181 LLSCLTTPASAYEVRNVSGIYHVTNDCSNSISIVYEAADVIMHTPCGVCVQEGNSSRCWV 240
Db 181 LLSCLTTPASAYEVRNVSGIYHVTNDCSNSISIVYEAADVIMHTPCGVCVQEGNSSRCWV 240
```

Db 181 LLSCLTVPASAYQVRNSSLGYHVTNDCPNSSIVYEADAILHTPCVPCVREGNASRCWV 240
Qy 241 ALTPTLAARNAASVPTTTRRHVDLVAGTAAPCSAMVVDLGCSSIFLVSQLPFTFSRRHET 300
Db 241 AVTPVATRDGKLPPTQURRHLDLVGSATLCSALYVGLDSCSVFLVQLFTFSRRHWT 300
Qy 301 VDCNCSIPYGHVSGHRMAWDMNMNWSPTTALVVSQLLRIPOAVVDMVAGAHGWGLAGLA 360
Db 301 TDCNCSIPYGHITGHRMAWDMNMNWSPTAALVVAQLLRIPOAIMDIAGAHGWGLAGIA 360
Qy 361 YYSVMGNNAKVLIVALLFAGVDGEHTTGRVAGHTTSFTSFPSSGASQKQLQVNTNGSW 420
Db 361 YFSVMGNNAKVLVLLFAGVDAETHVTGGSAGRTTAGVLGLLTPCAKQNIQLINTNGSW 420
Qy 421 HINRTALNCNDSLOTFPFAALFAVAKFNSSGCPERWASCRPIDWFAQCGWPTTYTKPNSS 480
Db 421 HINSTALNCNENLNTGWLTAFLYQHKFNSSGCPERLASCRRLTDFAOQGWPIISYANGSGL 480
Qy 481 DORPYCWHYAPRPCGVVPASQVCGPVYCFTPGPVVVGTTRDGSVPTYSGWNETDVMLIN 540
Db 481 DERPYCWHYPPRPCGLVPKSVCGPVYCFTPGPVVVGTTRDGSAPTYSGWANDTDVFLVN 540
Qy 541 NTRPPQGNWFGCTWMNSTGFTKTCGPPCNIIGVGNRTLICTDCFRKHPEATYKCSG 600
Db 541 NTRPPGNWFGCTWMNSTGFTKVCGAPPCVIGVGNNTLLCPTDCFRKHPEATYKCSG 600
Qy 601 PMLTPRCLVDYPRYLWHYCTINFISIFKVRMYGVGVEHRLNAACNWTGERCNLEDRDRS 660
Db 601 PWITPRCWDYPRYLWHYCTINYITFKVRMYGVGVEHRLNAACNWTGERCNLEDRDRS 660
Qy 661 ELSPLLSTTEWQILPACFTTLPALSTGLIHLHONIVDQYLYGVCSAFVSPFAIKWEYIL 720
Db 661 ELSPLLSTTQOVLPCTSTTLPALSTGLIHLHONIVDQYLYGVGSSSTASWAIKWEYV 720
Qy 721 LFLFLLLADARVACALMMLLLIAQEAALENLVLNAASVAGAHGILSLFLVFFCAAWYK 780
Db 721 LFLFLLLADARVCSLWMLLLISQEAALENLVLNAASLAGTHGLVSLVFFCFWAWYK 780
Qy 781 RIAPGAAYFYGVWPLLLLLLLALPPRAYALDREMAASCAGVLVGLVFTILSPYKVFELT 840
Db 781 RWVPGAAYFYGVWPLLLLLLLALPPRAYALDREMAASCAGVLVGLVFTILSPYKVFELT 840
Qy 841 RLILWLQYFTTRAEAHMWVPLNVGRGRDAIILLTCAVHPBELFDITKLLALLGLPLM 900
Db 841 WCMWWLQYFLTRVEAQHLWVPLNVGRGRDAVILLMVCVHPFLVFDIITKLLALLFGPLW 900
Qy 901 VLOAGITRVPYFYFRAOGLIRACMLRKVAGHYVQVFMKLGALITGYVYVNHILTPLRDWA 960
Db 901 ILOASLLKVPYFVRVOGLLRI CALARKIAGGHVQMAIILKLGALITGYVYVNHILTPLRDWA 960
Qy 961 HAGRLDAVAVEPVFVSAMETKITWGADTAACGDIILGLPVARSRGKEIFLGPADSLEG 1020
Db 961 HNGRLDAVAVEPVFVSAMETKITWGADTAACGDIINGLPVARSRGKEIFLGPADGMSV 1020
Qy 1021 OGWRLLAPITAYSOOTRGVLGCIITSLTGRDKNOVEGEVQVSTATQCSFLATCINGCVMT 1080
Db 1021 KGWRLLAPITAYAQOQRLGLGCIITSLTGRDKNOVEGEVQVSTATQOFLATCINGCVMT 1080
Qy 1081 VYHGAGSKTLAGPKPITQMYTNVDLDLVGWAQPPGARSMTPCSGSSDLYLVTRIADVI 1140
Db 1081 VYHGAGTRITIASPKGPVI QMYTNVDQDLVGNWPAQGSRLTPTCCGSSDLYLVTRIADVI 1140
Qy 1141 PVRRRGDSRGSLLSPRPVSYLKGSSGGLPLCPSGHVGVVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRRGDSRGSLLSPRPISYLKGSSGGLPLCPAGHAVGLFRAAVCTRGVAKAVDFIPVEN 1200
Qy 1201 METTMSPVTDNSTPAPVQOTQVAHLHAPTSKGSKTKVPAAYAAQGYKVLVLPNSVAA 1260
Db 1201 LETTMSPVTDNSTSPAPVQSFQVAHLHAPTSKGSKTKVPAAYAAQGYKVLVLPNSVAA 1260
Qy 1261 TLGFGAYMSKAHGDINRTIGVRTITTGSGITYSTYKFLADGCGSGGAYDIICDECHS 1320
Db 1261 TLGFGAYMSKAHGDVDPNRTIGVRTITTGSPITTYSTYKFLADGCGSGGAYDIICDECHS 1320

Qy 1321 TDSTTILGIGTVLDOAETAGARLVVLATATPPGSGVTVPHPNIEBIGLNNGEIPIFYGKAI 1380
Db 1321 TDATSILGIGTVLDOAETAGARLVVLATATPPGSGVTVPHPNIEBVALSTTGEIPIFYGKAI 1380
Qy 1381 PIEAIKGRHLIFCHSKKKCDELAALKTGLGLNAVAYRGLDVSVIPPIGVDVVVATDAL 1440
Db 1381 PLEVIKGRHLIFCHSKKKCDELAALKTGLGLNAVAYRGLDVSVIPITSGVDVVVSTDAL 1440
Qy 1441 MTGFTGPDSDSDICNTCVTQVDFSLDPTFTTETTTVPQDAVSRSSORGRGRSGIYR 1500
Db 1441 MTGFTGPDSDSDICNTCVTQVDFSLDPTFTTETTTVPQDAVSRQRRGRTGRGKGIYR 1500
Qy 1501 FVTGCRPSGMFDSVSLCECYDAGCAWYELTPAETSRLRAYLNTPLGVPQODHLEFWES 1560
Db 1501 FVAGCRPSGMFDSVSLCECYDAGCAWYELTPAETTVRLRAYMNTPLGVPQODHLEFWEG 1560
Qy 1561 VFTGLTHIDAHFLSOTKOAGDNFPYLVAQYATVCARAQAPPPSDWDMWKCLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSOTKSGENFPYLVAQYATVCARAQAPPPSDWDMWKCLIRLKPTLHG 1620
Qy 1621 PTPLLYRLGAVONEVILTHPITKYIMACMSADLEVTTSTWLVGVLAALAAAYCLITGCV 1680
Db 1621 PTPLLYRLGAVONEVILTHPITKYIMTMSADLEVTTSTWLVGVLAALAAAYCLUSTGCV 1680
Qy 1681 VIVGRIILSGKPAVVVPDREVLYQBFDEMEECASOLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Db 1681 VIVGRIILSGKPAIIPDREVLYQBFDEMEECASOLPYIEQGMQLAEQFKQKALGLLOTAT 1740
Qy 1741 KOEAAAPVWESKWRALSTFWAKHMNFIISGQIYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 ROAEVITPBAVQTNQKLEVPWAKHMNFIISGQIYLAGLSTLPGNPAIASLMAFTAAVTSP 1800
Qy 1801 LTTQNTLLFNLIGWVAQAOLAPPASAFVAGTAGAGVSGIGLQKVLVDIILAGVAGVA 1860
Db 1801 LTTQNTLLFNLIGWVAQAOLAAPGAATFVAGTAGAGVSGIGLQKVLVDIILAGVAGVA 1860
Qy 1861 GALVAFKVMSEVSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEAVOMNRLLI 1920
Db 1861 GALVAFKVMSEVSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEAVOMNRLLI 1920
Qy 1921 AFASRGNHVSPTHYVPESDAAARVTQILSSLTITQLLRLHQWINECSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHYVPESDAAARVTAILSSLTITQLLRLHQWISSECTTPCSGSWLRDI 1980
Qy 1981 WDWICTVLTDKFWLQSKLLPRLPGVPPSLCORGKGVWGGGIMQTTCPGCAQIAGHVK 2040
Db 1981 WDWICEVLSDFKTLWKAKLMPQLPGIPFVSCQGRYGVWGGGIMHTRCHCGAEITGHVK 2040
Qy 2041 NGSMEIIVGPRTCSNTWGHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVETRVGDPH 2100
Db 2041 NGTMRIVGPRTCRNWSGTFPINALTGCTPLPAPNYKFALRVSAEYVEIRRVGDPH 2100
Qy 2101 VYTGMTTNDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLLEDVTVFQVGLNQYLVGSOL 2160
Db 2101 YVSGMTTNDNLKPCQIPSEPEFFTELDGVRHLRHFAPCKPULLREEVSVFVGLHEYPVGSOL 2160
Qy 2161 PCEPEPDVTVLTSMLTDPSSHITAETAKRRLARGSPSPASSASQISAPSLKATCTTHHD 2220
Db 2161 PCEPEPDVAVLTSMLTDPSSHITAETAAAGRRRLARGSPSPASSASQISAPSLKATCTANHD 2220
Qy 2221 SPDAELIENLILWROEMGNITRVESENKVTILDSFEPLHABGDEREISVAEILKRSRK 2280
Db 2221 SPDAELIENLILWROEMGNITRVESENKVTILDSFDPLVAEEDREVSVPABILKRSR 2280
Qy 2281 FPSALPIWARPDYNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPPRKRRTVLTESNV 2340
Db 2281 FARALPVMARPDYNPPLVETWKPDYEPVHVHGCPLPPRSPPVPPPRKRRTVLTESNL 2340
Qy 2341 SSALAEALATKTPGSSGSAVSGTATALPDLASDDGDKGSDVESYSSMPLPEGPDPL 2400
Db 2341 STALAEALATKSPGSSSTSGITDNTTTSSEFAPGCPDSDVESYSSMPLPEGPDPL 2400

Qy	2401	SDGSWTSVSEEA-SEDVWCSSMYTWGALITPCAEEBSKLPINPLNSLLRHHNNVYAT	2459
Db	2401	SDGSWTSVSSGADTVVCCSSYSWTGALVTPCAAEEQKLPINALNSLLRHHNLVYST	2460
Qy	2460	TSBSASLRQKVTDFRLOVLDHYRDLVKEMKAKASTVVKAKLISIEEACKLTPPHSAKSK	2519
Db	2461	TSBSACORQKVTDFRLOVLDHSHYQDVLKEVKAAASKVXANLUSVBEACSLTPPHSAKSK	2520
Qy	2520	FGYGAIXDVRNLSRAVNHRSVWEDLEDTETPIDTTIMAKSEVFCVQPEKGGKRPALRI	2579
Db	2521	FGYGAIXDVRNLSRAVNHRSVWEDLEDTETPIDTTIMAKSEVFCVQPEKGGKRPALRI	2580
Qy	2580	VFPDLGVRCVKMALVDVSTLPOAVMGSSYGFQYSPKQKRVFELVNTWSKCKPMGFSYD	2639
Db	2581	VFPDLGVRCVKMALVDVSTLPOAVMGSSYGFQYSPKQKRVFELVNTWSKCKPMGFSYD	2640
Qy	2640	TRCFDSTVTSIDIRVESIYQCCDLAPAPQARQARSLSLTERLYIGGPLTNSKGQNGYRRCR	2699
Db	2641	TRCFDSTVTSIDIRTESAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSKGQNGYRRCR	2700
Qy	2700	ASGVLTTCSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLWVICESAGTQEDAAALRAFT	2759
Db	2701	ASGVLTTCSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLWVICESAGTQEDAAALRAFT	2760
Qy	2760	EMATRYSAAPPDPPQPEYDLELITSCSSNVSAHDASGKRVYLLTRDPTTPLARAAWETA	2819
Db	2761	EMATRYSAAPPDPPQPEYDLELITSCSSNVSAHDAGKRVYLLTRDPTTPLARAAWETA	2820
Qy	2820	RHTPINSWLGNIIMYATLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLP	2879
Db	2821	RHTPINSWLGNIIMFAPTLWARMILMTHFFSVLIARDQLEALNCEIYGACYSIEPLDLP	2880
Qy	2880	QIIRLHGLSAFTLHYSYSGEINRVASCLRKLGVPPPLTRWHRARSVRALKLSOGGGAAT	2939
Db	2881	PIQRLHGLSAFTLHYSYSGEINRVAAACLRKLGVPPPLTRWHRARSVRALKLSOGGGAAT	2940
Qy	2940	CGRYLFNVAVRTKLKLTPIPAASQDLDSGMFVAGYSGGDIYHLSRARPRWFLCLLLLS	2999
Db	2941	CGRYLFNVAVRTKLKLTPIPAASQDLDSGMFVAGYSGGDIYHLSRARPRWFLCLLLLS	3000
Qy	3000	VGVIYLLPNR 3010	
Db	3001	AGVGIYLLPNR 3011	
RESULT 7			
US-10-184-150-3			
; Sequence 3, Application US/10184150			
; Publication No. US20030134297A1			
; GENERAL INFORMATION:			
; APPLICANT: Olson, William			
; APPLICANT: Maddon, Paul			
; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS INFECTION			
; FILE REFERENCE: 2048/64896-A/JPM/MAF/DJK			
; CURRENT APPLICATION NUMBER: US/10/184,150			
; CURRENT FILING DATE: 2002-12-10			
; PRIOR APPLICATION NUMBER: 09/891,894			
; PRIOR FILING DATE: 2001-06-26			
; NUMBER OF SEQ ID NOS: 3			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 3			
; LENGTH: 3011			
; TYPE: PRN			
; ORGANISM: Hepatitis C virus			
US-10-184-150-3			
Query Match 88.3%; Score 14143.5; DB 4; Length 3011;			
Best Local Similarity 85.6%; Pred. No. 0;			
Matches 2576; Conservative 218; Mismatches 216; Indels 1; Gaps 1;			
Qy	1	MSTNPKPORKTKRNNRRPQDKVFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG	60
Db	1	MSTNPKPORKTKRNNRRPQDKVFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG	60

Qy	61	RRQIPKARPEGRAWAQPGYWPPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120
Db	61	RRQIPKARPEGRWAQPGYWPPLYGNEGCWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120
Qy	121	KVIDTLTCCFADLMGYIPLVGAAPLGGARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCCFADLMGYIPLVGAAPLGGARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA	180
Qy	181	LLSCLTIPASAEYVRNVSGIYHVNTDCSNSSIVYEADVIMHTPGCVPCVQVQENSRCWV	240
Db	181	LLSCLTIPASAEYVRNVSGIYHVNTDCSNSSIVYEADAILHTPGCVPCVQVQENSRCWV	240
Qy	241	ALPTTLAARNASVPTTIRRHVDLLVGTAAFCASAMVVDL-CGSI ELVSOLFESPRRHET	300
Db	241	AVTPTVATRDGKLPFTTQLRRHIDLLVGSATLCSALYVGDLCGSVFLVQGLFTFSPPRHWT	300
Qy	301	VQDCNCSIYPGHVSGHRMADMMNWSPTTALVVSQLLRIPQAVVDMVAGAHGVLGAGLA	360
Db	301	TQDCNCSIYPGHITGHRMADMMNWSPTAALVVAQLLRAIPQAIMDIAGAHGVLGAGLA	360
Qy	361	YYSMVGNWAKVLIVALLFAGVDGETHTRVAGHTTSGFTSLFSSGASQIKIQLVNTGWSW	420
Db	361	YFSMVGNWAKVLIVALLFAGVDAETHVTGGSAGRTTAGLVGLLTPGAKQNIQLINTGWSW	420
Qy	421	HINRTALNCNDSLOTGFPAALFYAHKFNSSGCCPERMASCRPIDWFAQGWGPITYTKPNSS	480
Db	421	HINSTALNCNDSLNTGWLGLFYQHKNFNSSGCCPERLASCRRLTDFAQGWGPISYANGSGL	480
Qy	481	DQRPYCHWYAPRPGVVPASQVCGPVYCTFTSPVAVGTTRDSGVPTYSWGENETDVMNLN	540
Db	481	DERPYCHWYAPRPGVVPASQVCGPVYCTFTSPVAVGTTRDSGVPTYSWGENETDVMNLN	540
Qy	541	NTPPQCNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLCTPTDCFRKHPEATYTKCGSG	600
Db	541	NTPPQCNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLCTPTDCFRKHPEATYTKCGSG	600
Qy	601	PWITPRCLVDYPRVRLMHPCTLNFSIPKVMYVGGVEHRLNAACNWTGRGERCNLEDRDS	660
Db	601	PWITPRCMVDYPRVRLMHPCTINVTIPKVMYVGGVEHRLNAACNWTGRGERCNLEDRDS	660
Qy	661	ELSPLLSTTQWQLPCSFLLTHLHQNIVDVOYLVGSGSAFVSFAIKWEYIL	720
Db	661	ELSPLLSTTQWQLPCSFLLTHLHQNIVDVOYLVGSGSIASWAIKWEYIV	720
Qy	721	LLFLLADARVCACLWMMLLIAQAEAALENLVINAASVAGAHGILSFLVFFCAAWYIKG	780
Db	721	LLFLLADARVCCCLWMMLLIISQAEAALENLVINAASLAGTHGLVSLVFFCFANWYIKG	780
Qy	781	RLAFGAAYAFYGVWPLLLLLLPPRAYALDREMAASCGGAVLVGLVFLTSLPYKVFILT	840
Db	781	RWVFGAVYAFYGVWPLLLLLLPPRAYALDREMAASCGGAVLVGLVFLTSLPYKVFILT	840
Qy	841	RLIHWLOFYITRAEHNQVWVPLNVGRDRAIILLCAVHPELIFDITKLLAILLGLPLM	900
Db	841	WCMWLOFYITRAEHNQVWVPLNVGRDRAIILLCAVHPELIFDITKLLAILLGLPLM	900
Qy	901	VLQAGITRVYFVRAOGLIRACMLVRKVAGHYVQVMFKLGTGTVVYVNHLPRLDWA	960
Db	901	ILQASLLKVPYFVRVQGLLRICALARKIAGHYVQMAIIKUGALTGTVVYVNHLPRLDWA	960
Qy	961	HAGRLDLAVAVEPVVFSAMETKVITWGADTAACGDIILGLFVSARRKEIFLGPADSLEG	1020
Db	961	HNGRLDLAVAVEPVVFSAMETKVITWGADTAACGDIILGLFVSARRKEIFLGPADSLEG	1020
Qy	1021	QGWELLAPITAYSQOTRGVLGCIITSLTGRDNQVGEVQVWSTATQSLATCINGUCWT	1080
Db	1021	KGWELLAPITAYSQOTRGVLGCIITSLTGRDNQVGEVQVWSTATQSLATCINGUCWT	1080
Qy	1081	VTHGAGSKTLAGKGPITQMTYTNVDLDLVGHQAPPGARSMTPCSCGSDLVIVTRHADVI	1140
Db	1081	VTHGAGTRTIAAPKGPVITQMTYTNVDLDLVGHQAPPGARSMTPCSCGSDLVIVTRHADVI	1140

1141 PVRRGDSRGSLLSPRPVSYLYKSSGGPILLCPSHGVGVFRAAVCTRGVAKAVDFIPVES 1200
Db PVRRGDSRGSLLSPRPISYLYKSSGGPILLCPAGHAVGLFRAAVCTRGVAKAVDFIPVEN 1200
1201 METTWRSPVFTDNSTPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQGVKVLVLPNSVAA 1260
Db METTWRSPVFTDNSTPPAVPQSFQVAHLHAPTGSKSTKVPAAAYAAQGVKVLVLPNSVAA 1260
1261 TLGFAGYMSKAGHDIPNIRGTGRTITGGSITYSTYKFLADGGSGGAYDIIICDECHS 1320
Db TLGFAGYMSKAGHDIPNIRGTGRTITGGSITYSTYKFLADGGSGGAYDIIICDECHS 1320
1321 TDSITILGIGTVLDOAETAGARLVVLATATPPGSGVTVPHPNIEEIGLSNNGSIPFYGKAI 1380
Db TDSITILGIGTVLDOAETAGARLVVLATATPPGSGVTVPHPNIEEVALSITGGEIPFYGKAI 1380
1381 PTEALIKGGRHLIFCHSKKKCDELAACL TGLGLNAVAYYRGLDVSVIPIGDVVVVVATDAL 1440
Db PLEVINGKGRHLIFCHSKKKCDELAACLVALGINAVAYYRGLDVSVIPTSGDVVVVYSTDAL 1440
1441 MTGFTGDFDSVLDKNTCVTQTVDFSLDPTFTTETTTVPDAYSRSORRGRTGRSGGIYR 1500
Db MTGFTGDFDSVLDKNTCVTQTVDFSLDPTFTTETTTLPDAYSRTQRRGRTGRGPGIYR 1500
1501 FVTPGERPSGMPDSSVLCEDYDAGCAWYELTPAETSVRLRAYLINTPGLPVCQDHLEFWES 1560
Db FVAPGERPSGMPDSSVLCEDYDAGCAWYELTPAETTVRLRAYLINTPGLPVCQDHLEFWEG 1560
1561 VFTGLTHIDAHFLSQTQKQAGNPPYLVAYQATVCARAQAAPPSPSWDMKCLIRLKPRTLHG 1620
Db VFTGLTHIDAHFLSQTQKQGENFPYLVAYQATVCARAQAAPPSPSWDMKCLIRLKPRTLHG 1620
1621 PTELLYRLGAVONEVTLTHPIKYTKMACNSADLEVVTSTVWLGVGLAALAAAYCLTTGCV 1680
Db PTELLYRLGAVONEVTLTHPIKYTKMCSADLEVVTSTVWLGVGLAALAAAYCLSTGCV 1680
1681 VTVGRILSGKPAVDPREVLVQEFDEMEECASQLPIEQGMOLAEQFKQKALGLLOTAT 1740
Db VIVGRILVSGKPAIPDREVLVQEFDEMEECQHLPIYEQGMMLAEQFKQKALGLLOTAS 1740
1741 KQAEAAAPVVEKSWRALETFWAKHMNFISGIIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db ROAEVITPAVQTNWQLEVFWAKHMNFISGIIQYLAGLSTLPGNPAIASLMAFTAAVTSP 1800
1801 LTTQNTLLNILLGGWAAOLA PPSASAFVGGAGIAGAAVSGTIGKVLVDIILAGYGAGVA 1860
Db LTTQNTLLNILLGGWAAOLAAPGAATAFVAGLAGAAITGSGVLGKVLVDIILAGYGAGVA 1860
1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGAVQWNNRLI 1920
Db GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGAVQWNNRLI 1920
1921 AFASRGNHVSPTHYVPESDAARVTOILSSLTITQLLKLHWINBEDCTPCSGSWLRDV 1980
Db AFASRGNHVSPTHYVPESDAARVTAISSLTVTLRLRLHQLWISSECTTPCPSGSLRDI 1980
1981 WDWICTVLTDFKTLWOSKLLPRLPGVPFLSCQRYGVWRGDMQTTCPGCAOIAGHVK 2040
Db WDWICEVLSDFKTLWAKLMPQLPGIPFVSCQRYGVWRGDMHTRCHCAEITGHVK 2040
2041 NGSMTVGPRTCSNTWHGTFPINAYTTPCTSPAPNYSRALMRVAAEYVEVTRVGDPH 2100
Db NGTMRIVGPRTCNMMSGTFPINAYTTPCTPLPAPNYKFAALWRVSAEYVEIRVGDPH 2100
2101 YVTGMTDNVKKCPQVPAPEFFTEVDGVLRLHRYAPACKPLLRBDVTTFQVGLNOYLVGSOL 2160
Db YVSGMTDNLKCPQIPSPSEFFTELDGVLRLHRYAPPCKPLLRBEVFSRVLGHEYPVGSOL 2160
2161 PCEPEPDVTLTSMLTDPGSHITAEAKRLARGSPPSLSSASQLSAPSLKATCTTHHD 2220
Db PCEPEPDVTLTSMLTDPGSHITAEAGRLARGSPPSMASSASQLSAPSLKATCTTANHD 2220
2221 SPDADLIEANLWRQMGGNITRVESENKVILDSPEPLHAEGDEREISVAAEILKRSRK 2280

2221 SPDAELIEANLWRQMGGNITRVESENKVILDSFDPLVAEEDEREVSVAEILKRSR 2280
QY FPSALPTWARDPYNPPLLESKWKDPYVPPVVGCPCLPTKAPPJPPRRKRTVVLTEGNSV 2340
Db FARALPVWARDYNPPLVETWKKPDYEPVVGCPCLPPRPPPPPPRKRRTVVLTESTL 2340
2341 SSALAEALATKTFGSSGSAVDSGTATAPLDLASDDGDKGSDVESYSSMPPLEGEFGDPL 2400
Db STALAEALATKTFGSSSTSGITGDNNTTTSSEFAPSGCPCPDSDVESYSSMPPLEGEFGDPL 2400
2401 SDGWSVTVEEA-SEDEVCCSMTVTGALTTPCAABESKLPINPLSNLSLRHHNMVAT 2459
Db SDGWSVTSSGADTDEVCCSMTSWTGCALVTPCAABEQKLPINALSNLSLRHHNLVST 2460
2460 TSSASLRQKVTDRLOVLDHYRDLKEMKASTVKAKLSTEEACKLTPPHSAKS 2519
Db TSSACORQKVTDRLOVLDHYRDLKEMKASTVKAKLSTEEACKLTPPHSAKS 2520
2520 FGAKADVRNLSRAVNHIRSWEDLLEDTPIDTTIMAKSEVFCVQPEKGRKPARLI 2579
Db FGAKADVRCHARVAHINSWKDLLEDVTPIDTTIMAKNEVFCVQPEKGRKPARLI 2580
2580 VFPDLGVRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQRVEFLVNTWKSCKPMGFPSYD 2639
Db VFPDLGVRVCEKMALYDVVSKLPLAVMGSSYGFQYSPQRVEFLVQAWKSCKTPMGFSYD 2640
2640 TRCFDSTVTESDIRVEESIYOCCLAPEARQAIKSLTERLYIGGLPLTNSKGCNCCYRRCR 2699
Db TRCFDSTVTESDIRTEEAIOCCDLDPQARVAIKSLTERLYIGGLPLTNSRGENCCYRRCR 2700
2700 ASGVLTTSCGNLTCLYKATACRAAKLQDCTMLVNGDDLVIICBSAGTQEDAAALRAFT 2759
Db ASGVLTTSCGNLTCLYKARACRAAGLQDCTMLVCGDDLVIICBSAGVQEDAAALRAFT 2760
2760 EAMTRYAPPDPPQPEYDLELITSCSNVSAVHADAGKRVYVLTDRPTTPLARAAMETA 2819
Db EAMTRYAPPDPPQPEYDLELITSCSNVSAVHADAGKRVYVLTDRPTTPLARAAMETA 2820
2820 RHTPINSWLGNIIMVAPTLWARMILMTHFFSILLAQEOLKALDQOICACYSIEPLDLP 2879
Db RHTPINSWLGNIIMVAPTLWARMILMTHFFSILLAQEOLKALDQOICACYSIEPLDLP 2880
2880 OIIEHLGLSAPFTLHSPGIBNRVASCRLKLGVPPLTRWHRARSVRAKLISQCGRAAT 2939
Db OIIEHLGLSAPFTLHSPGIBNRVASCRLKLGVPPLTRWHRARSVRARULLSRGAAI 2940
2940 CGRYLFNMAVRTKLTLPAAASQDLDSGWVAGYSGGDIYHSLSRARPRWFPLCLLLLS 2999
Db CGRYLFNMAVRTKLTLPAAAGRLDLSGWFTAGYSGGDIYHVS HARPRWFPLCLLLIA 3000
3000 VGVGIYLLPNR 3010
3001 AGVGIYLLPNR 3011

RESULT 8

US-10-328-997-3
; Sequence 3, Application US/10328997
; Publication No. US20030232745A1
; GENERAL INFORMATION:
; APPLICANT: Olson, William
; APPLICANT: Maddon, Paul
; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS INFI
; FILE REFERENCE: 2048/64896-B/JPW/MAF/DJK
; CURRENT APPLICATION NUMBER: US/10/328,997
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: 09/891,894
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 10/184,150
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3																										
; LENGTH: 3011																										
; TYPE: PR																										
; ORGANISM: Hepatitis C virus																										
US-10-328-997-3																										
Query Match																										
Best Local Similarity 88.3%; Score 14143.5; DB 4; Length 3011;																										
Matches 2576; Conservative 218; Mismatches 216; Indels 1; Gaps 1;																										
QY	1	MSTNPKPQRTKNTNR	PDVAFPGGGQIVGGVILLPRG	PLGYRATRKASERSQPRG	60	QY	961	HAGLRDLAVAVPEPVFSAMETKVIITW	GADTAACGDIILGLPVSARRGKEIFLGP	ADSLRG	1020															
DB	1	MSTNPKPQRTKNTNR	PDVAFPGGGQIVGGVILLPRG	PLGYRATRKATSERQPRG	60	DB	961	HNGLRDLAVAVPEPVFSRMETKLIITW	GADTAACGDIILGLPVSARRGQIILGP	ADGMVS	1020															
QY	61	RROPIPKARRPEGRAWAQ	PGYPWPLVYCNELGWAGLLSPRG	SRPWSGPTDPRRRSRNLG	120	QY	1021	QWRLLAPITAYSQOTR	GVLCIIITSLTGRDNQKQVEGEVQV	STATQSF	LATCINGVCWT	1080														
DB	61	RROPIPKARRPEGRAWAQ	PGYPWPLVYCNELGWAGLLSPRG	SRPWSGPTDPRRRSRNLG	120	DB	1021	KWRLLAPITAYSQOTR	GLGCIITSLTGRDNQKQVEGEVQI	VSTATQT	OTFLATCINGVCWT	1080														
QY	121	KVIDTLTCGFADLMGYI	PLVGAFLGGAARALAHGVR	VLEDGVN	YATGNLPGCSFSIFLLA	180	QY	1081	VYHGAGSKTLAPKGP	ITQMYTNVDLVLVGWQAPPGAR	SMTPTCS	CGSSDLYL	YTRHADVI	1140												
DB	121	KVIDTLTCGFADLMGYI	PLVGAFLGGAARALAHGVR	VLEDGVN	YATGNLPGCSFSIFLLA	180	DB	1081	VYHGAGTRTITASPK	GPVIQMYTNVDQDLVGM	PAQGSRLT	PCTCGSSDLYL	YTRHADVI	1140												
QY	181	LSCLTIPASAYVRNVSGI	YHYHTNDCNSSI	YVEAADVIMHT	PGCVPCVOEGNSRCWV	240	QY	1141	PVRRGDSGSLSRP	PVSYLKGSSGPLL	CSGHVVGVFRAAVCT	TRGVAKAVDFI	PVES	1200												
DB	181	LSCLTIPASAYQVRN	SSGLYHTNDCPNSSI	YVEAADAILHT	PGCVPCVREGNASCWV	240	DB	1141	PVRRGDSGSLSRP	ISYLGSSGGPLC	PAGHAVGLFRAAVCT	TRGVAKAVDFI	PVEN	1200												
QY	241	ALTPTLAARNASVPTT	TIRBHVLLVGTAAFC	SAMVVDLCGSI	FLVSQLFTFSPPRHET	300	QY	1201	METTMRSVPTDNST	PPAVPQTFQVAHLHAPT	SGSKSTKVPAA	YAAAGQKVL	LNPSVAA	1260												
DB	241	AVTPTVATROGKUPT	QOLRRHIDLLVGSAT	LCALYVGDLCGS	FVLVGQLFTFSPPRHWT	300	DB	1201	LETTHMRSVPTDN	SSPPAVPQSFQVAHLHAPT	SGSKSTKVPAA	YAAAGQKVL	LNPSVAA	1260												
QY	301	VQDCNCSIYPGHVSGH	RAMDMNMNSPTTAL	VVSQLLRIPOAV	DMVAGAHGVLAGLA	360	QY	1261	TLGFGAYMSKAHGI	DPNIRTVRTIT	TGGSITVSTY	GKFLADGCGSGGAYDII	ICDECHS	1320												
DB	301	TQDCNCSIYPGHITG	HREAMDMMNSPTAAL	VVAQLLRIPQAIMD	MIAGAHGVLAGIA	360	DB	1261	TLGFGAYMSKAHGV	DPNIRTVRTIT	TGGSPIITYSTY	GKFLADGCGSGGAYDII	ICDECHS	1320												
QY	361	YYSWGVNNAKVLIV	ALLFAGVDGTHET	TRVAGHTTSGF	TSLSFGSGASQIK	LVNTGWSW	420	QY	1321	TSTTILGIGTVLDQ	AEAGARLVVLATAT	PPGCVTVPHPNIE	IEIGLSNNGEI	PFYCKAI	1380											
DB	361	YFSWGVNNAKVLV	LLLVLLFAGVDAETH	VTGGSAGRTTAGL	VGLLTPGAKQI	LIQINTGWSW	420	DB	1321	TDATSIIGIGTVLDQ	AEAGARLVVLATAT	PPGCVTVSHPNIE	BEVALSTTGEI	PFYCKAI	1380											
QY	421	HINRTALNCNDSLOT	GFEALFYAHKFNSS	CGPERMASC	RPIWFAQGWGP	ITVTKPNSS	480	QY	1381	PIBAIKGGRHILF	CHSKKCKDELA	AKLIGLGN	AVAYYRGLDVS	PIPPIGD	VVVVATDAL	1440										
DB	421	HINSTALNCNESLT	GWLAGLFYQHKFNSS	CGPERLASC	RRLTDFAQGWGP	ISTANGSGL	480	DB	1381	PLEVIKGGRHILF	CHSKKCKDELA	AKLIGLGN	AVAYYRGLDVS	PIPTSGD	VVVVATDAL	1440										
QY	481	DORPCYCHYAPR	CGVVPASQVCPV	CFYTPSPVVVGT	DRSGVPTY	SWGENETD	VMLLN	QY	1441	MTGFTGDFSD	VIDNCVTQVDF	SLDPTFTI	ETTTVPD	AVRSQR	RGTGRSGIYR	1500										
DB	481	DERPCYCHYAPR	CGIYPAKSVCPV	CFYTPSPVVVGT	DRSGAPTY	SWGANDID	VFLN	DB	1441	MTGFTGDFSD	VIDNCVTQVDF	SLDPTFTI	ETTTLPQ	DASVRSQR	RGTGRSGIYR	1500										
QY	541	NTRPPQGNWFGCT	WMNSTGFTKTCG	PPNCNIGGVN	RTLICPTDCFR	KHPBATY	KCGSG	QY	1501	FVTPGERPSQ	MFSSVLC	ECYDAGCAW	YELTPAETS	VRLRAYLNT	PGLPVCODH	LEWES	1560									
DB	541	NTRPPLGNWFGCT	WMNSTGFTKVC	GAPPCVIGGVGN	TLLCPTDCFR	KHPBATY	SRCGSG	DB	1501	FVAPGERPSQ	MFSSVLC	ECYDAGCAW	YELTPAETS	VRLRAYMNT	PGLPVCODH	LEWEG	1560									
QY	601	PWLTPRCLVDY	VPRLMHPY	CTLNFSIFKRV	MYVGVGVEHRL	NAACNWT	RGERCNLEDR	DRS	QY	1561	VFTGLTHID	DAFLISQTKQ	SGENFPYL	VAYQATV	CARAQAPPP	SDQMKCLIRL	KPTLHG	1620								
DB	601	PWITPRCMVDY	VPRLMHPY	CTINVTIFKRV	MYVGVGVEHRL	EAAACNWT	RGERCDLEDR	DRS	DB	1561	VFTGLTHID	DAFLISQTKQ	SGENFPYL	VAYQATV	CARAQAPPP	SDQMKCLIRL	KPTLHG	1620								
QY	661	ELSPILLSTTEWQIL	PCAF	TTLPALSTGLIHL	QNI	VDQVLYGV	GSFAVSFAIKWEVIL	720	QY	1621	PTLLYRLG	AVQNEVIL	THPIKYIM	ACMSAD	LEVVTSTWV	LGVLAALAA	YCLTTCGV	1680								
DB	661	ELSPILLSTTQWVLP	CSFTTLPALSTGLIHL	QNI	VDQVLYGV	GSFAVSFAIKWEVYV	720	DB	1621	PTLLYRLG	AVQNEVIL	THPIKYIM	ACMSAD	LEVVTSTWV	LGVLAALAA	YCLTTCGV	1680									
QY	721	LLFLLADARV	CACILMMMLL	IAQAEALLEN	VNNAASV	AGHGLS	FLVFFCAAWYIKG	780	QY	1681	VIVGRIL	ILSGKPAV	VPDREVL	YQFDE	MEBCASQ	LPVIEQGM	QIAE	QFKALGLLQ	TAT	1740						
DB	721	LLFLLADARV	CSCLMMMLL	ISQAEALLEN	VILNNAAS	LAGTH	GLVSLVFFCF	AWYIKG	DB	1681	VIVGRIV	ILSGKPAI	IPDREVL	YQFDE	MEBCS	QHLPIVIEQGM	QIAE	QFKALGLLQ	TAS	1740						
QY	781	RLAPGAAV	AFYGVW	PLLLALLAL	PPRAYAL	DREMAAS	CGGAVL	GLVFLTIS	PPYKVELT	840	QY	1741	KQAEAAA	PVVESK	WRALET	FWAKMWNFI	SGIQV	LAGLSTL	PGNPATAS	LMFA	TASITSP	1800				
DB	781	RWPGAV	AVFYGVW	PLLLALLAL	PQRAV	ALDTEVAAS	CGGVVL	VGLMAL	TLSPPYK	YIS	QY	1741	QAEVIT	PAVQTN	WQKLE	VFVWAKMWNFI	SGIQV	LAGLSTL	PGNPATAS	LMFA	TAAVTS	1800				
QY	841	RLIIMW	LOVITRAE	AHMQW	VPPLN	VRGRDAI	ILLITCAV	PELLIFD	ITKLLAL	IGLM	QY	1801	LTQNT	LLFNIL	LGWVAQ	LAAPSA	SAFV	CAGI	AGAAV	SGIGL	KVLVD	ILAGY	GAGVA	1860		
DB	841	WCMWML	LOVITRAE	QALH	WVPLN	VRGRDAV	ILLIMC	VVHPT	LFVDF	ITKLLAL	IGLM	QY	1801	LTQNT	LLFNIL	LGWVAQ	LAAPSA	SAFV	CAGI	AGAAV	SGIGL	KVLVD	ILAGY	GAGVA	1860	
QY	901	VLOAGIT	RVPEV	RAQGLIR	ACMLVR	KVAGGHV	QVWMFK	GALGT	YVYNH	LTPLRDWA	960	QY	1861	GALVAF	KVMS	GEVSTED	VNLLP	AILSPG	ALV	VGVVCA	AILRRH	VGP	EGAVQ	MMNRLI	1920	
DB	901	ILQASLL	KVPEV	RVYR	VOGLLRI	CALARKI	AGGHVQ	MAIKL	GALGT	YVYNH	LTPLRDWA	960	DB	1861	GALVAF	KVMS	GEVSTED	VNLLP	AILSPG	ALV	VGVVCA	AILRRH	VGP	EGAVQ	MMNRLI	1920

Qy	2041	NGSMRIVGPRPTCSNTHGTPPIINAYTGPCTPSPADNYSBALMRVAAAEVVEVTRVGDFH	2100
Db	2041	NGTMIRIVGPRPTCNMMSGTFPIINAYTGPCTPLPAPNYKEFALMRVSAEYVEIRRVGDFH	2100
Qy	2101	YVGTMTTNDNVKCPQVPAPEFFTEVDGVRHLRHVAPACKPLLREDVTFOVGLNOYLVGSOL	2160
Db	2101	YVSGMTTNDNKKPCQIIPSPBEFFTELDGVRUHRFAPCKPLUREVSVFRUGLHYBPVGSOL	2160
Qy	2161	PCPEPDVTVLTSMLTDPSPHITAFKAKRRILARGSPPSLASSASSQSAPLSKATCTTHHD	2220
Db	2161	PCPEPDVAVLTSMLTDPSPHITAEAAAGRRILARGSPPSMASSASSQSAPLSKATCATNHD	2220
Qy	2221	SPADLIEANLLWRQEMGNIITRVESENKVILDSFEPLHABGDREISVAAAILRKSRK	2280
Db	2221	SPDAELIEANLLWRQEMGNIITRVESENKVILDSFDPLVAEEDEREVSVPAILRKSR	2280
Qy	2281	FPSALPIWARPDYNNPPLLESWKOPDYVPVVVHGCPLPPTKAPIPPPRRKRTTVVLTESNV	2340
Db	2281	FARALPWARPDPYNNPLVETWKPEDYEPVVVHGCPLPPPSPPVPPPRKRTTVVLTESTL	2340
Qy	2341	SSALAEIATKTFGSGSSAVDSGTATALPDLASDDGDKGSDVSYSYSGMPPLEGEPGDPDL	2400
Db	2341	STALAEIATKSGFSSSTSGITGDNNTTTSSEPAPEGCPDSDVESYSYSGMPPLEGEPGDPDL	2400
Qy	2401	SDGSWSTVSEEA-SEDEVCCSMSTVTGALITPCAABESKLPIINPLNSLLRHHNMVYAT	2459
Db	2401	SDGSWSTVSGGATEDVVCSCMSYSTGALVTPCAABEQKLPINALNSLLRHHNLVYST	2460
Qy	2460	TSRSASLRQKVTFDRLQVLDDHYRDLVKEMKAKASTVKAKLASIEACKLTPPHSAKSK	2519
Db	2461	TSRSACORQKVTFDRLQVLDSHVQDVLKEVKAASVKANLLSVEACSLTPPHSAKSK	2520
Qy	2520	FGYGAQDVRNLSRAVNHIIRSWBEDLLEDETETPIDTTIMAKSEVFCVQPKGGRKPARLI	2579
Db	2521	FGYGAQDVRCHARKAVAHINSVMKDLLEDSVTPIDTTIMAKNEVFCVQPKGGRKPARLI	2580
Qy	2580	VFPDLGVRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQRQVEFLVNTWKSKCPMGFSYD	2639
Db	2581	VFPDLGVRVCEKMALYDVVSKPLVAVMGSSYGFQYSPQRQVEFLVQAWKSKTIMGFSYD	2640
Qy	2640	TRCFDSTVTESDIRVEESIYQCQDLAPEARQAIRSLTERLYIGGPLTNSXGQNGYRRCR	2699
Db	2641	TRCFDSTVTESDIRTEAIYQCQDLDPQARVAIKSLTERLYVGGPLTNSRGNGCYRRCR	2700
Qy	2700	ASGVLTSTSCGNLTLCYLKATAACBAAKLODCTMLVNGDGLVWICESAGTOEDAAALRAFT	2759
Db	2701	ASGVLTSTSCGNLTLCYIKARAACRAAGLODCTMLVCGDDLVIICESAGVQEDAAALRAFT	2760
Qy	2760	EMATRYGAPDPPQPEYDLELITSCSSNVSAHDAGKRVYVYLTRODPTTPLARAAMETA	2819
Db	2761	EMATRYGAPDPPQPEYDLELITSCSSNVSAHDAGKRVYVYLTRODPTTPLARAAMETA	2820
Qy	2820	RHTPINSWLGNIIMYAPTLWARMILMTHFTSILLABQLEKALDCQIYGACYSIEPDLDP	2879
Db	2821	RHTPVNSWLGNIIMFAPTLWARMILMTHFTSVLIARQLEQALNCEIYGACYSIEPDLDP	2880
Qy	2880	QIIRLHGLSAPTLHSSPGEINRVASCLRKLGVPPLRTWHRARSVRAKLISOGGRAAT	2939
Db	2881	PIQRLHGLSAPTLHSSPGEINRVAAACLRLKLGVPPLRAWHRARSVRALLSRGGRAAI	2940
Qy	2940	CGRYLFNWAVRTKLLKLTPIPAASOLDLSGMFWAGYSGGDIYHLSLSRARPRWFFPLCLLLLS	2999
Db	2941	CGKYLFWNAVRTKLLKLTPIAAAGRLDLSGMFTAGYSGGDIYHVSYSHARPRWFFWFLCLLLLA	3000
Qy	3000	VGVGIYLLPNR 3010	
Db	3001	AGVGIYLLPNR 3011	

```

1  GENERAL INFORMATION:
2  APPLICANT:  RICE, CHARLES  et al.
3  TITLE OF INVENTION:  FUNCTIONAL DNA CLONE FOR HEPATITIS C
4  TITLE OF INVENTION:  VIRUS (HCV) AND USES THEREOF
5  NUMBER OF SEQUENCES:  21
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
8  STREET:  7733 FORSYTH BLVD., SUITE 1400
9  CITY:  ST. LOUIS
10 STATE:  MO
11 COUNTRY:  USA
12 ZIP:  63105
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE:  Floppy disk
15 COMPUTER:  IBM PC compatible
16 OPERATING SYSTEM:  PC-DOS/MS-DOS
17 SOFTWARE:  Patent In Release #1.0, Version #1.30
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER:  US/09/238,076
20 FILING DATE:
21 CLASSIFICATION:
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER:  US 09/034,756
24 FILING DATE:
25 ATTORNEY/AGENT INFORMATION:
26 NAME:  HOLLAND, DONALD R.
27 REGISTRATION NUMBER:  35,197
28 REFERENCE/DOCKET NUMBER:  6029-4831
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE:  314-727-5188
31 TELEFAX:  314-727-6092
32 INFORMATION FOR SEQ ID NO:  2:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH:  3012 amino acids
35 TYPE:  amino acid
36 STRANDEDNESS:  single
37 TOPOLOGY:  linear
38 MOLECULE TYPE:  protein
39 HYPOTHETICAL:  NO
40 FRAGMENT TYPE:  N-terminal
41 US-09-238-076-2

```

RESULT 9

US-09-238-076-2

US-09-238-076-2
: Sequence 2: Application US/09238076

; Patent No. US20020102540A1

[illegible]

Qy	1501	FVTPGPRPSGMFSSVLCBCYDAGCAGWYELT	PAETISVRLRAYLNT	PGLPV	CDODHLEF	WES	1560
Db	1501	FVAPGPRPSGMFSSVLCBCYDAGCAGWYELT	PAETITVRLRAYLNT	PGLPV	CDODHLEF	WEG	1560
Qy	1561	VFTGLTHIDAHFLSQTQKAGDNPPYLVA	QATVCARAQAPPPSD	QMMKCLIRL	KPTLHG	1620	
Db	1561	VFTGLTHIDAHFLSQTQKAGDNPPYLVA	QATVCARAQAPPPSD	QMMKCLIRL	KPTLHG	1620	
Qy	1621	PTPLLRLGAVQNEVILTHPI	TKYIMACHMSADLEV	TSWLV	VGGVLAALAY	CLTTGSV	1680
Db	1621	PTPLLRLGAVQNEVILTHPI	TKYIMACHMSADLEV	TSWLV	VGGVLAALAY	CLTTGSV	1680
Qy	1681	VIVGRITILSKCPAVDPREVL	XOEFDEMEECASQ	LPYIEQGMOLABOF	KOKALGLLQ	TAT	1740
Db	1681	VIVGRITILSKCPAVDPREVL	XOEFDEMEECASQ	LPYIEQGMOLABOF	KOKALGLLQ	TAT	1740
Qy	1741	KQEAAPVYVESKWRALFTFWAKHMWNF	ISGIOYLAGLSTL	PCGNPAIASI	MAFTASITSP	1800	
Db	1741	ROAEVITPAVQTNWQKLEVPWAKHMWNF	ISGIOYLAGLSTL	PCGNPAIASI	MAFTAAVTSP	1800	
Qy	1801	LTTQNTLLFNILGWWAAQ	LAPPSAASAFV	GAGIAGAAVGS	IGLKVLDI	LAGYGVA	1860
Db	1801	LTTQNTLLFNILGWWAAQ	LAAPCAATAFV	GAGLAGAAIGS	VGLGKVLVDI	LAGYGVA	1860
Qy	1861	GALVAFKVMGSEVPSTEDL	VNLPAILSPCAL	VVGVCALIRRHV	PGCEGAVOMNRLI	1920	
Db	1861	GALVAFKIMGEVPSTEDL	VNLPAILSPCAL	VVGVCALIRRHV	PGCEGAVOMNRLI	1920	
Qy	1921	AFASRGNHVSPTHVVPESDAAA	RVTOILSSLTIT	QLLKRLHOMIN	EDCSTPCSGS	MLRDV	1980
Db	1921	AFASRGNHVSPTHVVPESDAAA	RVTOILSSLTIT	QLLKRLHOMIN	EDCSTPCSGS	MLRDV	1980
Qy	1981	WDWICTVLTDFKWLQSKL	PLRPLPGVPFL	SCQRQYKGVMRG	DGIMOTTC	PCGAQIAGHVK	2040
Db	1981	WDWICEVLSDFKWLKAKL	MPQLPGIPFV	SCQRQYGRVWEG	DGIMHTRCHCGAE	ITGHVK	2040
Qy	2041	NGSMRIVGPRTCSNTHG	TPIINAYTTG	PTCPSBPAPNY	SPALNRVAAE	VEVETR	2100
Db	2041	NGTWRIVGPRTCRNMW	SGTPIINAYTTG	PTCPAPNYPK	PALNRVAAE	VEVETR	2100
Qy	2101	YVTGWTITDNVUKPCOV	PAPPEFTE	DGVRHLRVAPACK	PLLRDVTFOV	GLNOYLVGSOL	2160
Db	2101	YVSGMTITDNLKCPCQ	IPSPBFFTE	DGVRHLRVAPACK	PLLRREVSPR	GLHYPVGSOL	2160
Qy	2161	PCEPEPDVLTLSMLT	DPSHITAETAKRRL	ARGSPPSLASSAS	QLSAPSLKATCT	THHD	2220
Db	2161	PCEPEPDVAVLTLSMLT	DPSHITAEAGRL	ARGSPPSMASSSS	QLSAPSLKATC	TANH	2220
Qy	2221	SPDADLIEANLLWRQ	MGNIITRVES	ENKVVILDSPEPL	HAEGDERIS	VAAEILRKSRK	2280
Db	2221	SPDAELIEANLLWRQ	MGNIITRVES	ENKVVILDSFPL	VAEEDERE	VSVPAEILRKSR	2280
Qy	2281	FPSALPIWARDPYNP	PLLESKWD	PDYVPVPHG	CPPLPTKAPPI	PPRRKRTVVL	2340
Db	2281	FARALPVMARDPYNP	PLVETWKKPDYEP	VPHGCPPLPR	SPVPVPRKRTVVL	TESTL	2340
Qy	2341	SSALAEATKTFGSGS	ANDSGAT	ALPDLASD	GDGKSDVES	YSWMPLEGE	2400
Db	2341	STALAEATKTFGSGS	TSITGDN	TTSSPEAP	SGCPTD	SDVSYSNMPLEGE	2400
Qy	2401	SDGWSMTVSEEA-	SEDVCCMSY	TWTGALIT	PCAABESKL	PINPLNSL	2459
Db	2401	SDGWSMTVSSGAD	TEDVCCMSYS	WTGALVTP	CAABEQKL	PINALNSL	2460
Qy	2460	TSRSASLSQKVT	FDRLQVLD	DDHYRDL	KEMKAAST	VKAKLIS	2519
Db	2461	TSRSACQKQKVT	FDRLQVLD	SHYQDVL	KEVKAAS	VKANKLL	2520
Qy	2520	FGYGAKDVRNLS	SSRAVNHIRS	WEDLLED	TETPTD	TTIMAKSE	2579
Db	2521	FGYGAKDVRCHARK	AINSWKDL	LED	SVTP	DDTTIMAKNE	2580
Qy	2580	VFPDILGVVCFKMA	LYDVVSTL	POAUMGSS	YGFQYSPKOR	VEFLVNTW	2639

Db 2581 VFPDLGVRVCERWALYDVVSKPLFLAVMGSSYGFQISPGQORVEFLVQAWKSKTKPMGFSYD 2640
Qy 2640 TRCFDSTVTESDIRVEESIYOCCLAPARQAIRSITELRYLIGGPLTNSKQONCCYRRCR 2699
Db 2641 TRCFDSTVTESDIRTEEAIIYQCCDLDPQARVAIKSITELRYLIGGPLTNSRGENCYRRCR 2700
Qy 2700 ASGVLTTCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVIICBSAGTQEDAAALRAFT 2759
Db 2701 ASGVLTTCGNTLTCYLKATAACRAAGLQDCTMLVCGDDLVIICBSAGVQEDAAALRAFT 2760
Qy 2760 EAMTRYSAPPPOPEYDLELITSCSNVSVAHADSGKRVYLLTRDPTTPLARAAMETA 2819
Db 2761 EAMTRYSAPPPOPEYDLELITSCSNVSVAHADSGKRVYLLTRDPTTPLARAAMETA 2820
Qy 2820 RHTPNSWLGNIIMVAPTILWARMILMTHEFFSILLAQEQLEKALDQOIGACYSIEPLDLP 2879
Db 2821 RHTPNSWLGNIIMVAPTILWARMILMTHEFFSILLAQEQLEKALDQOIGACYSIEPLDLP 2880
Qy 2880 QIIRLHGLSAPTLHSYSPEINRVASCLKLGVPPLRTWRHRARSVRAKLLSCGGRAT 2939
Db 2881 PIIRLHGLSAPTLHSYSPEINRVASCLKLGVPPLRTWRHRARSVRAKLLSCGGRAT 2940
Qy 2940 CGRYLFNMAVRKLTPIPAASQDLSCGFVAGYSGGDIYHSLRARPWFPLCLLLLS 2999
Db 2941 CGRYLFNMAVRKLTPIPAASQDLSCGFVAGYSGGDIYHSLRARPWFPLCLLLLS 3000
Qy 3000 VGVGIYLLPNR 3010
Db 3001 AGVGIYLLPNR 3011

RESULT 10

US-09-995-937-2
; Sequence 2, Application US/09995937
; Publication No. US20030028010A1
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 28-Nov-93
; APPLICATION NUMBER: US/09/995,937
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,756
; FILING DATE: 04-May-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3012 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-995-937-2

Query Match 88.3%; Score 14143.5; DB 3; Length 3012;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 2576; Conservative 218; Mismatches 216; Indels 1; Gaps 1;

Qy 1 MSTNPKPQKTKRNTNRRPQDVKFPGGQIVGGVYLLPRGRPLGVRAKRSERSQRG 60
Db 1 MSTNPKPQKTKRNTNRRPQDVKFPGGQIVGGVYLLPRGRPLGVRAKRSERSQRG 60

Qy 61 RRQPIPKARRPEGRAWAOPGYWPPLYGNEGGLGAGWLLSPGRSRPSWGPTDPRRSRNIG 120
Db 61 RRQPIPKARRPEGRAWAOPGYWPPLYGNEGGLGAGWLLSPGRSRPSWGPTDPRRSRNIG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRLVEDGVNYATGNLPGCSFIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRLVEDGVNYATGNLPGCSFIFLLA 180

Qy 181 LLSCLTTPASAYEVNRVNSGIYHVYVNDSCNSSIIVYBAADVIMHTPCVPCVQEGNSRCWV 240
Db 181 LLSCLTTPASAYEVNRVNSGIYHVYVNDSCNSSIIVYBAADVIMHTPCVPCVQEGNSRCWV 240

Qy 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYVGLDGLSGIIFLVSQLFTFSPRRHET 300
Db 241 AVTPTVATRDGKLPPTQRRHIDLLVGSATLCSAIYVGLDGLSVFLVQQLFTFSPRRHET 300

Qy 301 VQDCNCSIYPGHVSGHRMAWMMNWSPTTALVWSQLLRIPOAVVDMVAGAHGWVLGAGLA 360
Db 301 VQDCNCSIYPGHVSGHRMAWMMNWSPTTALVWSQLLRIPOAVVDMVAGAHGWVLGAGLA 360

Qy 361 YYSVGNWAKVLIIVALLPAGVDGETHTTGRVAGHTTSGTSLFSSGASQKIQLVNTGNSW 420
Db 361 YYSVGNWAKVLIIVALLPAGVDGETHTTGRVAGHTTSGTSLFSSGASQKIQLVNTGNSW 420

Qy 421 HINRTALNCDSLOTGFFAALPYAHKFNSSGCPERMASCRPIDWFAQCGMPTTYTKPNSS 480
Db 421 HINRTALNCDSLOTGFFAALPYAHKFNSSGCPERMASCRPIDWFAQCGMPTTYTKPNSS 480

Qy 481 DORPYCMHYAPRCPGVWPASQVCGPVYCFPTSPVYVGTDRSGVPTYSWGENETDVMNLN 540
Db 481 DORPYCMHYAPRCPGVWPASQVCGPVYCFPTSPVYVGTDRSGVPTYSWGENETDVMNLN 540

Qy 541 NTRPPQGNWFGCTWNNSTGFTKTCGGPPCNTGGVGNRTLICPTDCFRKHPEATYTKCGSG 600
Db 541 NTRPPQGNWFGCTWNNSTGFTKTCGGPPCNTGGVGNRTLICPTDCFRKHPEATYTKCGSG 600

Qy 601 PHLTPRCLVDYPYRLWHYPCTLNFSI FKVRYMVGVGVEHRLNAACTWRCERCNLEDRDS 660
Db 601 PHLTPRCLVDYPYRLWHYPCTLNFSI FKVRYMVGVGVEHRLNAACTWRCERCNLEDRDS 660

Qy 661 ELSPILLSTTEWQIILPCAFITLTPALSTGLIHLHQNIQVQYLYGVGSFAFVSKWEYIL 720
Db 661 ELSPILLSTTEWQIILPCAFITLTPALSTGLIHLHQNIQVQYLYGVGSFAFVSKWEYIL 720

Qy 721 LFLLLADARVCACLMWMLLIAQAEALLENVLVNAASVAGAHGILSLFLVFFCAAWYIKG 780
Db 721 LFLLLADARVCACLMWMLLIAQAEALLENVLVNAASVAGAHGILSLFLVFFCAAWYIKG 780

Qy 781 RLAPGAAYAFYGVWPILLILLALPRAYALDREMAASCGGAVLVGLVLTLSPPYKVELT 840
Db 781 RLAPGAAYAFYGVWPILLILLALPRAYALDREMAASCGGAVLVGLVLTLSPPYKVELT 840

Qy 841 RLIMWLQYFPIRAEAHMOWVPLNVGRGDRAIILLTCAVHPLELIDITKLLAILGLPM 900
Db 841 RLIMWLQYFPIRAEAHMOWVPLNVGRGDRAIILLTCAVHPLELIDITKLLAILGLPM 900

Qy 901 VLOAGITRVPYFVRAQGLIRACMLVRKAGHYVQVFMKLGALITGYTYNHLTPLRDWA 960
Db 901 VLOAGITRVPYFVRAQGLIRACMLVRKAGHYVQVFMKLGALITGYTYNHLTPLRDWA 960

QY 961 HAGLRDLAVAVEPVFSAMETKVIITWADTAACGDIILGLPVARSARGKEIFLGPADSLG 1020
DB HNGLRDLAVAVEPVFSAMETKVIITWADTAACGDIILGLPVARSARGKEIFLGPADSLG 1020
QY 1021 OGRHLLAPITAYSOOTRGVLGCIITSITGRDKNQVEGVVSTQATSLATCINGVCWT 1080
DB KGRHLLAPITAYSOOTRGVLGCIITSITGRDKNQVEGVVSTQATSLATCINGVCWT 1080
QY 1081 VHGAGSKTLGKPGPITOMYTNVDLVLGQAPPGARSMTPCSCGSSDLIYLVRHADVI 1140
DB VHGAGSKTLGKPGPITOMYTNVDLVLGQAPPGARSMTPCSCGSSDLIYLVRHADVI 1140
QY 1141 PVRRRGDSRGLSLPRPVSYLKGSSGGPLLCPSGHVGVVPRAAVTRGVAKAVDFIPVES 1200
DB PVRRRGDSRGLSLPRPVSYLKGSSGGPLLCPSGHVGVVPRAAVTRGVAKAVDFIPVES 1200
QY 1201 METTMRSPVFTDNSTPAPVQOTQVAHLHAPTCGSKTKVPAAYAAQYKVLVLPNSVAA 1260
DB LETTMRSPVFTDNSTPAPVQOTQVAHLHAPTCGSKTKVPAAYAAQYKVLVLPNSVAA 1260
QY 1261 TLGFGAYMSKAHGDPIRGTGRTITGSGITTYSTYKFLADGCGSGAYDIIICDECHS 1320
DB TLGFGAYMSKAHGDPIRGTGRTITGSGITTYSTYKFLADGCGSGAYDIIICDECHS 1320
QY 1321 TDSITLIGIGTVLDQAETAGARLVLATATPPGSVTVPHNPIEIGLSNNGEIPFYGKAI 1380
DB TDATSLIGIGTVLDQAETAGARLVLATATPPGSVTVPHNPIEIGLSNNGEIPFYGKAI 1380
QY 1381 PIERAIGGRHLIFCHSKKCKDELAUKTGLGNNAVYRGDLVSVIPIGDVVVVATDAL 1440
DB PLEVIKGRHLIFCHSKKCKDELAUKTGLGNNAVYRGDLVSVIPIGDVVVVATDAL 1440
QY 1441 MTGFTGDFSDVIDCNTCTOTVDFSLDPTFTIETTVPODAVSRORRGTRGRGIYR 1500
DB MTGFTGDFSDVIDCNTCTOTVDFSLDPTFTIETTVPODAVSRORRGTRGRGIYR 1500
QY 1501 FVTPGERPSGMFSSVLCECYDAGCAMEYLTPTAETSVRLRAYLNTPLGPVQCDHLEFWS 1560
DB FVAPGERPSGMFSSVLCECYDAGCAMEYLTPTAETSVRLRAYLNTPLGPVQCDHLEFWS 1560
QY 1561 VFTGLTHIDHFLSQTQAGDNFPYLVAYQATCARAQAPPPSDQWKKLIRLKTPLHG 1620
DB VFTGLTHIDHFLSQTQAGDNFPYLVAYQATCARAQAPPPSDQWKKLIRLKTPLHG 1620
QY 1621 PTPLLYRLGAVQNEVILTHPTIKYIMACMSADLEVTSWLVGCVLAAALAYCLTTGVS 1680
DB PTPLLYRLGAVQNEVILTHPTIKYIMACMSADLEVTSWLVGCVLAAALAYCLTTGVS 1680
QY 1681 VIVGRIILSGKPAVVPDREVLYQBFDEMEECASQLPYIEQGMQLAEQFKQKALGLLOTAT 1740
DB VIVGRIILSGKPAVVPDREVLYQBFDEMEECASQLPYIEQGMQLAEQFKQKALGLLOTAT 1740
QY 1741 KOABAAAPVBSKWRALFTFWKHMWNPISIGIYVLAGLSTLPGNPAIASLMAFTASITSP 1800
DB ROAEVITPAVQNTWQLEWVFKHMWNPISIGIYVLAGLSTLPGNPAIASLMAFTASITSP 1800
QY 1801 LTTQNTLLFNILGGVAAQLAPPSAASAFVGNAGTAGAVSGIGLKVLDVILAGVAGVA 1860
DB LTTQNTLLFNILGGVAAQLAPPSAASAFVGNAGTAGAVSGIGLKVLDVILAGVAGVA 1860
QY 1861 GALVAFKMSGEVSPSTEDVNLPAILSPGALVGVVCAALIRRHVGPGEAVQWMMRLI 1920
DB GALVAFKMSGEVSPSTEDVNLPAILSPGALVGVVCAALIRRHVGPGEAVQWMMRLI 1920
QY 1921 AFASRGNHVSPTHYVSPESDAARVTQIILSSITITQLLKRLHQWINEDECSTPCSGSLRDV 1980
DB AFASRGNHVSPTHYVSPESDAARVTQIILSSITITQLLKRLHQWINEDECSTPCSGSLRDV 1980
QY 1981 WDWICTVLTDKFTWLOSKLLPRLPGVDFLSCORGKGVGRGDGIMQITTCPCGAQIAGHVK 2040
DB WDWICEVLSDFKTLWKALMPQLPGIPFVSCORGKGVGRGDGIMHTRCHCGAEITGHVK 2040

RESULT 11

US-09-917-563-2

; Sequence 2, Application US/09917563

; Publication No. US20030073080A1

QY 2041 NGSMRIVGPRCTSNTHGTFPINAYTTGCTPSPAPNYSRALWRVAAEEVVEVTRVGDFFH 2100
DB NGTMIRIVGPRCTCRNMWSGTFPINAYTTGCTPLPAPNYPALWRVSAEYVEIRRVGDFFH 2100
QY 2101 YVTGMTDNNVCPQCPVAPPEFFTEVDGVRHLRYAPACKPLLRDDVTFQVGLNQYLVGSOL 2160
DB YVSGMTDNLKCPQCPVAPPEFFTEVDGVRHLRYAPACKPLLRDDVTFQVGLNQYLVGSOL 2160
QY 2161 PCEPEPDVTVLSMLTDPSSHITAETAKRRLARGSPSSLASSASOLSAPELKATCTTHHD 2220
DB PCEPEPDVAVLSMLTDPSSHITAETAKRRLARGSPSSLASSASOLSAPELKATCTTHHD 2220
QY 2221 SPDAULIEANLLWRQSMGNIITRVESENKVILDSPEPLHAEGDEREISVAAAILKRSRK 2280
DB SPDAULIEANLLWRQSMGNIITRVESENKVILDSPEPLHAEGDEREISVPAILKRSRR 2280
QY 2281 FFSALPIWARPDYNNPILLESWKDPDYVPVHVHCCPLPPTKAPPIPPRRKRTVVLTESNV 2340
DB FARALPVWARPDYNNPILLESWKDPDYVPVHVHCCPLPPTKAPPIPPRRKRTVVLTESNV 2340
QY 2341 SSALAEIATKTFGSSGSSAVDSGTATAPLADSDGDKGSDVESYSMPLEGEPCGPDLL 2400
DB STALAEIATKTFGSSGSSAVDSGTATAPLADSDGDKGSDVESYSMPLEGEPCGPDLL 2400
QY 2401 SDGSWSTVSEEA-SEDVVCSSMSYTWGALITPCAABESKLPIINPLSNLLRHHNMVYAT 2459
DB SDGSWSTVSSGADTEDVVCSSMSYTWGALITPCAABESKLPIINPLSNLLRHHNMVYAT 2459
QY 2460 TSRSASIRQKVTFDRLQVLDDHYRDLVKEMKAKASTVAKALLSIEBACKLTPPHSAKSK 2519
DB TSRSACORQKVTFDRLQVLDDHYRDLVKEMKAKASTVAKALLSIEBACKLTPPHSAKSK 2519
QY 2520 FGYGAKDVRLSSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVFCVQPEKGGKPARLI 2579
DB FGYGAKDVRLSSRAVNHRSVWEDLLEDTEPTIDTTIMAKSEVFCVQPEKGGKPARLI 2579
QY 2580 VFPDLGVRCCKMALVDVNSTLPAQVWSSYGYQSPQVFEFLVNTWKSCKCPMGFSYD 2639
DB VFPDLGVRCCKMALVDVNSTLPAQVWSSYGYQSPQVFEFLVNTWKSCKCPMGFSYD 2639
QY 2640 TRCFDSVTESDIRVEBSIYQCCDLAPPEARQAIRSLTERLYIGGLTNSKQNGCYRRCR 2699
DB TRCFDSVTESDIRVEBSIYQCCDLAPPEARQAIRSLTERLYIGGLTNSKQNGCYRRCR 2699
QY 2700 ASGVLTTSCTNTLTCYUKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDDAALRAFT 2759
DB ASGVLTTSCTNTLTCYUKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDDAALRAFT 2759
QY 2760 EAMTRYSAPGDPDPQPEYDLELITSCSSNVSVAHDAAGKRVYVLTROPTTTLARAAMETA 2819
DB EAMTRYSAPGDPDPQPEYDLELITSCSSNVSVAHDAAGKRVYVLTROPTTTLARAAMETA 2819
QY 2820 RHTPINSWLGNIIMVAPTLWARMLTMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLP 2879
DB RHTPINSWLGNIIMVAPTLWARMLTMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLP 2879
QY 2880 QIIBERLHGLSAFTLHYSYSGEINRVASCLRKGLGVPLRTWHRARSVRKLLSOGGRAAT 2939
DB QIIBERLHGLSAFTLHYSYSGEINRVASCLRKGLGVPLRTWHRARSVRKLLSOGGRAAT 2939
QY 2940 CGRYLFNVAVRTKULPTIPAASQDLDSGFVAGYSGGDIYHSLSRARPRPFCLLLLS 2999
DB CGRYLFNVAVRTKULPTIPAASQDLDSGFVAGYSGGDIYHSLSRARPRPFCLLLLS 2999
QY 3000 VGVGIVLLPNR 3010
DB AGVGIYLLPNR 3011

GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSER: HOWELL & HAERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/917,563
FILING DATE: 27-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/238,076
FILING DATE: 26-JAN-1999
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3012 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-917-563-2

Query Match 88.3%; Score 14143.5; DB 3; Length 3012;
Best Local Similarity 85.6%; Pred. No. 0;
Matches 2576; Conservative 218; Mismatches 216; Indels 1; Gaps 1;
QY 1 MSTNPKPQKTKRNTNRRPDQVKFGGGQIVGGVYLLPRRGRPLGVRAITRKASERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRRPDQVKFGGGQIVGGVYLLPRRGRPLGVRAITRKASERSQPRG 60
QY 61 RRQPIKARRPREGAWAQGYWPLYGNGELGWAGHLLSPRGRSRPSWGPTDPRRNRNLG 120
DB 61 RRQPIKARRPREGRTWAQGYWPLYGNGCGWAGHLLSPRGRSRPSWGPTDPRRNRNLG 120
QY 121 KVIDTLTTCGFADLMGYIPLVGAPLGAARALAHGVRLDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTTCGFADLMGYIPLVGAPLGAARALAHGVRLDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIPASAEYVRNVSGIYHVTNDCNSSIYVEAADVIMHTPGVCVPCQEGNSRCRW 240
DB 181 LLSCLTIPASAEYVRNVSGIYHVTNDCNSSIYVEAADAILHTPGVCVPCVREGNASRCW 240
QY 241 ALPTLAARNASVPTTIRRHVDLLVGTAAFCSAMTVGLCGSIFLVSQLFTFSPRRHET 300
DB 241 AVTPTVATRDGKLPTQLRRHIDLVLGSAITLCSALYVGLCGSVFLVGQLFTFSPRRHWT 300
QY 301 VQDCNCSIYPGHVSGHRMADMMNWSPTTALVVSOLLRIPOAVDMVAGAHGWLAGIA 360
DB 301 VQDCNCSIYPGHITGHRMADMMNWSPTAALVVAOLLRIPOAIMDIIAGAHGWLAGIA 360
QY 361 YYSVMGNWAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW 420
DB 361 YYSVMGNWAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW 420

DB 361 YYSVMGNWAKVLIVALLFAGVDAETHVTGGSAGRTTAGLVGLLTPGAKQNIQIINTNGSW 420
QY 421 HINRTALNCNDLSLOTGFPFAALFYAHKFNSSCPERMASCRIPTDFWFAQMGPIITYTKPNSS 480
DB 421 HINSTALNCNESLNTGWLAGFYQHKFNSSCPERLASCRRTDFAQMGPIISYANGSGL 480
QY 481 DQPYCWHYAPRPGCVVPASQVCGPVYCTFSPVVVGTTRDSGVPTYSGENETDVMLN 540
DB 481 DERPYCWHYPPRPGCVVPASQVCGPVYCTFSPVVVGTTRDSGAPTYSGANDTDVFLN 540
QY 541 NTRPPQGNWFGCTWNSTGFTKTCGGPPCNTGGVGNRTLI CPTDCFRKHPEATYTKCGSG 600
DB 541 NTRPPQGNWFGCTWNSTGFTKTCGGPPCNTGGVGNRTLI CPTDCFRKHPEATYTKCGSG 600
QY 601 PWTPRCLVDYPRYLWHPYCTLNFSI FKVRYMVGVEHRLNAACNWTGERCNLEDRDS 660
DB 601 PWITPRCMVDYPRYLWHPYCTLNFSI FKVRYMVGVEHRLNAACNWTGERCNLEDRDS 660
QY 661 ELSPLLLSTTEWQILPCAFITLTPALSTGLIHLHQNIVDVQVLYGVGSAPVSPAIKWEYIL 720
DB 661 ELSPLLLSTTQWQVLPSCFTTLPALSTGLIHLHQNIVDVQVLYGVGSSTIASWAIKWEYV 720
QY 721 LFLLLADARVCACIMMMLLIAQAEALLENVLNAAVAGAHGILSFLVFFCAWYIKG 780
DB 721 LFLLLADARVCSCIMMMLLIQAEALLENVLNAAVAGAHGILSFLVFFCFAYILKG 780
QY 781 RLAPCAAYAFYGMVPLLLALLPRAVALDREMAASCGGAVLGLVFLTSLSPYKVFIL 840
DB 781 RWPGAVYAFYGMVPLLLALLPRAVALDREMAASCGGAVLGLVFLTSLSPYKRYIS 840
QY 841 RLIIWMLQYFITRAEAHMVWPVPLNVRGRDAIILLTCAVHPELIDITKLLAILGLPLM 900
DB 841 WCMWMLQYFLTEVEAQLHVWPVPLNVRGRDAVILLMCVHPTLVFDITKLLAILGLPLM 900
QY 901 VLQAGITRVYFVRAQGLIRACMLVRKAGHYVQVPMKLGALGTGYVYNNHLTPLDWA 960
DB 901 ILQASLLKVPYFVRVQGLLRICALARKIAGHYVQVPMKLGALGTGYVYNNHLTPLDWA 960
QY 961 HAGLRDLAVAVEPVVFSAMETKVTIWTGADTACGDI ILGLPVASARRGEIPLGADSLSG 1020
DB 961 HNGLRDLAVAVEPVVFSAMETKVTIWTGADTACGDI INGLPVASARRGEIPLGADGWS 1020
QY 1021 QWRLLAPITAYSQOTRGVLGCIITSLTRDKNQVEGEVQVYSTATQSFATCINGVCWT 1080
DB 1021 KGWRLAPITAYSQOTRGVLGCIITSLTRDKNQVEGEVQVYSTATQIFLATCINGVCWT 1080
QY 1081 VYHAGSKTLAGPKGPIQMYTNVDLVLVQWAPPGARSMTPCSCGSSDLYLVTRHADVI 1140
DB 1081 VYHAGSKTLAGPKGPIQMYTNVDLVLVQWAPPGARSMTPCSCGSSDLYLVTRHADVI 1140
QY 1141 PVRRGDSRGSILSPRPVSYLKGSSGGLPCPSGHVGVFRAAVCTRGVAKAVDFTPVES 1200
DB 1141 PVRRGDSRGSILSPRPVSYLKGSSGGLPCPSGHVGVFRAAVCTRGVAKAVDFTPVES 1200
QY 1201 METTMRSPFTDNTSPPAVPOTFQVAHLHAPTGSGSKTKVPAAYAAQGYKVLVLPNSVAA 1260
DB 1201 LETTMRSPFTDNTSPPAVPQSFQVAHLHAPTGSGSKTKVPAAYAAQGYKVLVLPNSVAA 1260
QY 1261 TLGFGAYMSKAHGDIPNIRITGVTITGGSITYSYGKPLADGGCGGGAYDIIICDECHS 1320
DB 1261 TLGFGAYMSKAHGDIPNIRITGVTITGGSITYSYGKPLADGGCGGGAYDIIICDECHS 1320
QY 1321 TDSITILGTVDQAEATAGARLVLATATPGSVTVPHNIEEGLSNNGEIPFVGKAI 1380
DB 1321 TDSITILGTVDQAEATAGARLVLATATPGSVTVPHNIEEGLSNNGEIPFVGKAI 1380
QY 1381 PTEALKGRGHLIFCHSKKKDELAALKGLGNVAVYRGLDVSVIPTGDDVVVATDAL 1440
DB 1381 PLEVILKGRGHLIFCHSKKKDELAALKGLGNVAVYRGLDVSVIPTGDDVVVATDAL 1440
QY 1441 MTGFTGDFSDVSDCNVCTQVDFSLDPTFTTITTTTTPQDAVRSORRORTGRSGIYR 1500
DB 1441 MTGFTGDFSDVSDCNVCTQVDFSLDPTFTTITTTTTPQDAVRSORRORTGRSGIYR 1500

QY 1501 FVTPGERPSGMPDSSVLCYCEYDAGCAWYELTTPAETSRLRAYLNTPLGPVCDHLEFWES 1560
DB 1501 FVAPGERPSGMPDSSVLCYCEYDAGCAWYELTTPAETTVRLRAYMNTFGLPVCQDHLFEWEG 1560
QY 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAVQATVCARAQAAPPSPSDQMWKCLRLKPTLHG 1620
DB 1561 VFTGLTHIDAHFLSQTQKAGDNFPYLVAVQATVCARAQAAPPSPSDQMWKCLRLKPTLHG 1620
QY 1621 PTPPLVRLGAVQNEVTLTHPIKYINMACSADLEVVTSTWLVGGVLAALAYCLTGTGV 1680
DB 1621 PTPPLVRLGAVQNEVTLTHPIKYINMACSADLEVVTSTWLVGGVLAALAYCLTGTGV 1680
QY 1681 VTVGRILISGKPAVDPREVLYQEFDEMBECASQLEPIBQGMQARQFKQKALGLIQTAT 1740
DB 1681 VIVGRIVLSGKPAIIPREVLYQEFDEMBECASQLEPIBQGMQARQFKQKALGLIQTAT 1740
QY 1741 KQAEAAAPVVEKSWRALETFWAKHMNFISGIQYLAGLSTLPONPAIASLMAFTASITSP 1800
DB 1741 ROAEVITPAVOTNWQKLEVFQWAKHMNFISGIQYLAGLSTLPONPAIASLMAFTAAVTSIP 1800
QY 1801 LTTONTLLFNILGCVAAQLAPSAASAFVAGIAGAAVSGIGLKVLDVILAGYAGVA 1860
DB 1801 LTTGQTLLFNILGCVAAQLAPSAASAFVAGIAGAAVSGIGLKVLDVILAGYAGVA 1860
QY 1861 GALVAFKMSGEVPESTEDLVNLLPAILSPGALVGVVCAILRRHVGPGEVAVQWNRLLI 1920
DB 1861 GALVAFKMSGEVPESTEDLVNLLPAILSPGALVGVVCAILRRHVGPGEVAVQWNRLLI 1920
QY 1921 AFASRGNHVSPTHYVPESDAAAARTQILSLTITQLLKRLHQMINEDECSPPCSGSLRDV 1980
DB 1921 AFASRGNHVSPTHYVPESDAAAARTQILSLTITQLLKRLHQMINEDECSPPCSGSLRDV 1980
QY 1981 WDWICTVLTDFKTLWOSKLLPRFLPGVFPFLSCQRYGKGVWRGDGIMQOTPCGQAQIAGHVK 2040
DB 1981 WDWICEVLSDFKTLWAKLMPGLPIFPVSCQRYGKGVWRGDGIMHTRCHCGAEITGHVK 2040
QY 2041 NGSMRIVGPRTCSNTHGTFPIINAYTTCPTSPAPNYSALWRVAABEVEVTRVGDFH 2100
DB 2041 NGTMRIVGPRTCNMWSGTFPIINAYTTCPTSPAPNYPKALWRVSAEYVEIRRVGDFH 2100
QY 2101 YVTGMTDNDKPCQVAPPEFFTEVDGRLHRRVAPACKPLLRREDVTQVGLNQVYLSQL 2160
DB 2101 YVSGMTDNDKPCQIPSPPEFFTELDGRLHRRVAPACKPLLRREVSFRVLGHEYPVGSOL 2160
QY 2161 PCPEPDVTVLTSMLTDPSSHITAEAKRLARGSPPSLASSASOLSAPSLKATCTTHHD 2220
DB 2161 PCPEPDVAVLTSMLTDPSSHITAEAGRRRLARGSPPSMASSSASOLSAPSLKATCTANHD 2220
QY 2221 SPDALEIANLLWRQEMGNITRVESENKVILDSFEPLEHAEGDEREISVAAEILRSRK 2280
DB 2221 SPDALEIANLLWRQEMGNITRVESENKVILDSFDPLEVAEEDEREVSVPABEILRSRR 2280
QY 2281 FPSALPIWARPDYNNPILLESWKDPDYPPVHVGCPPLPTKAPPTPPRRKRTVVLTSNV 2340
DB 2281 FARALPWARDYNNPILVETWKKEDYPPVHVGCPPLPPRSPVPVPRKXRTVVLTESTL 2340
QY 2341 SSALAEIATKTFGSGSADVSDGTATLPDLASDDGKSGDVSYSMPPLLEGPGPDDL 2400
DB 2341 STALAEIATKTFGSSSTSGITGNDNTTSSPEAPSGCPCPDVSYSMPPLLEGPGPDDL 2400
QY 2401 SDGSWSTVSEEA-SEDVVCCSMYSYTWGALTTPCAABESKLPIPNLSLRLHNNVYAT 2459
DB 2401 SDGSWSTVSSGADTEDVVCCSMYSYTWGALTTPCAABEQKLPINALSLSLRLHNNLVYST 2460
QY 2460 TSSASLRQKVVTFDRLQVLDHVDYDLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSK 2519
DB 2461 TSSACRQKVVTFDRLQVLSHVDYDLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSK 2520
QY 2520 FGYGAKDVRLSSRAVNHRSWEDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLI 2579
DB 2521 FGYGAKDVRCARKAVAHINSVKDLLEDSVTPIDTTIMAKNEVFCVQPEKGRKPARLI 2580

QY 2580 VPPDLGVRVCEKALYDVVSTLPOAVMGSSYGFQYSPKQORVEFLVNTWKSKECPMGFSYD 2639
DB 2581 VFPDLGVRVCEKALYDVVSKLPLAVMGSSYGFQYSPQORVEFLVQAWKSKTTPMGFSYD 2640
QY 2640 TRCFDSTVTESDIRVEESIYQCCDLAPEARQAIRSILTERLYIGGPIITNSKQNGCYRRRCR 2699
DB 2641 TRCFDSTVTESDIRVEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCK 2700
QY 2700 ASGVLTTCGNTLTCTVYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAFT 2759
DB 2701 ASGVLTTCGNTLTCTVIKARAAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAAASURAF 2760
QY 2760 EAMTRYSAAPPQPPQPEYDLELITSCSSNVSAHDASGKRVVYLTEDPTTLPARAAWETA 2819
DB 2761 EAMTRYSAAPPQPPQPEYDLELITSCSSNVSAHDGAGKRVVYLTEDPTTLPARAAWETA 2820
QY 2820 RHTPNSWLGNTIMVAPTLMWARMILMTHFFSILLAOEALDQCOIYGACYSIEPLDLP 2879
DB 2821 RHTPNSWLGNTIMFAPTLMWARMILMTHFFSVLIARDQLEALNCEIYGACYSIEPLDLP 2880
QY 2880 QIIRLHLGLSAFTLHSPGEINRVASCLRLKGLVPPPLRTWRHRARSVRAKLLSQGGRAAT 2939
DB 2881 PIILQRLHLGLSAFTLHSPGEINRVAAACLRKGLVPPPLRAWRHARSVRARLLSRGCGRAAI 2940
QY 2940 CGRYLPNMAVRTKLKLTPIPAASQDLSCGFWAGYSGGDIYHSLSRARPRFWPFLCLLLLS 2999
DB 2941 CGKYPNMAVRTKLKLTPIAAAGRLDLSGFWTAGYSGGDIYHSHVSHARPRFWPFLCLLLLA 3000
QY 3000 VGVGIYLLPNR 3010
DB 3001 AGVGIYLLPNR 3011

RESULT 12
US-10-296-734-406
; Sequence 406, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 406
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: HepC 1a consensus polyprotein
US-10-296-734-406

Query Match 88.3%; Score 14131.5; DB 4; Length 3011;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 2574; Conservative 216; Mismatches 220; Indels 1; Gaps 1;

QY 1 MSTNPKPQRTKRTNRRPQDVKPPGGQIVGGVYLLPRRGPRIGVTRATRKASERSOPRG 60
DB 1 MSTNPKPQRTKRTNRRPQDVKPPGGQIVGGVYLLPRRGPRIGVTRATRKTSERSOPRG 60
QY 61 RRQIPKARRPEGRWAPQPGYPWPLYNCEGLWAGWLLSPRSGRYSNGPTDPRRRSNL 120
DB 61 RRQIPKARRPEGRWAPQPGYPWPLYNCEGLWAGWLLSPRSGRYSNGPTDPRRRSNL 120
QY 121 KVIDTLTCGADLMGYIPLVGAAPLGAARALAHGVRLVDGVNYATGNLPCCSFSIFLLA 180
DB 121 KVIDTLTCGADLMGYIPLVGAAPLGAARALAHGVRLVDGVNYATGNLPCCSFSIFLLA 180
QY 181 LLSCLTTPASAYEVRNVSGIYHVNTDCSNSSIVAAADVIHMTPGCVPCVQEGNSRCW 240

[illegible]

1261	TLGFGAYMSKAHGIDPNIRITGVRTITTTGSPITTYTSTYGVKFLADGCGSGGAYDIIICDECHS	1326
1321	TDSITLIGIGTVLQDOASTAGARLVVLATATPPGSGVTVPHPNIEEIGLSNNGEIPFFYKAI	1386
1321	TDATSIILIGIGTVLQDOASTAGARLVVLATATPPGSGVTVPHPNIEEVALSITTEIPFFYKAI	1380
1381	PIEAIKGRGRLHIFCHSKKKCDELAALKLTGLGLNAVAYYRGLDVSVPIPGDVVVVATDAL	1440
1381	PLEVIKGRHLLIFCHSKKKCDELAALKVALGINAVAYYRGLDVSVIPTSGDVVVVATDAL	1440
1441	MTGTFGDFSVIDCNTCVTQTVDFSLDPTTETITTTVPQDAVRSQRGRGTRGRSGIYR	1500
1441	MTGYTGFDFSVIDCNTCVTQTVDFSLDPTTETITTTLPQDAVRSQRRGTRGKPGIYR	1500
1501	PVTGEPSPGWFSSVLCECVDAGCAWYELTPAETSVRLAYLNTPLPCVDOHLEPWES	1560
1501	FVAGEPSPGWFSSVLCECVDAGCAWYELTPAETTVRLRAYMNTPLPGVCQDHLEPWEG	1560
1561	VFTGLTHIDAHFLSOTQKAGDNFPYLVAQYATVCARAQAPPPSWDQWMMKCLIRLKPTLHG	1620
1561	VFTGLTHIDAHFLSOTQKSGENFPYLVAQYATVCARAQAPPPSWDQWMMKCLIRLKPTLHG	1620
1621	PTPLLRYLGAQVNEVILTHPIITYKIMACMSADLEVTVTWLVGGVLAALAAAYCLTTGSV	1680
1621	PTPLLRYLGAQVNEVILTHPVTKYINTCMSADLEVTVTWLVGGVLAALAAAYCLSTGCV	1680
1681	VIVGRIILSGKPAVVPDEVLVYOEFDMECSAQOLPYIEQGMOLAEQFKOKALGLQATAT	1740
1681	VIVGRIVLSGKPAIIPREVLRYREFDMECSOHLPYIEQGMMLAEQFKOKALGLLQATAS	1740
1741	KQAEAAAAPVVESEKWRALLETFWAKHMNFISGIOVLAGLSTLPGNPAIASLMAFTASITSP	1800
1741	QAEVIAPAVQTNWQKLEVPFWAKHMNFISGIOVLAGLSTLPGNPAIASLMAFTAATVSP	1800
1801	LTQTNTLLFNILGQWAAQALAPSAASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVA	1860
1801	LTTSTQTLFNILGQWAAQLAAPGAATAFVGAGLAGAAIGSVGLGKVLVDILAGYGAGVA	1860
1861	GALVAFKVMGSEVPSTEDLVNLLPAIILSPCALVVGUVVCAAILRRHHVGPGEVQVWNRLLI	1920
1861	GALVAFKIMGSEVPSTEDLVNLLPAIILSPCALVVGUVVCAAILRRHHVGPGEVQVWNRLLI	1920
1921	AFASRGNHVSPTHVVPESDAAARVTOILSSLTITQLLKRLHOMINEDCSTPFCSGSWLRDV	1980
1921	AFASRGNHVSPTHVVPESDAAARVTAIILSSLTITQLLRRLHQWISSECTTPFCSGSWLRDI	1980
1981	WDMTCTVLTDFTKLWQSKLAPRLPGVPFPLSCQRYGKGVWRGDGIMQTTCPGQAQIAGHVK	2040
1981	WDMICEVLSDFKWLKAKMLPQLPGIPFVFCQRYGKGVWRGDGIMHTRCHCGAEITGHVK	2040
2041	NGSMRIIVGPRTCNTWHTGTFPINAYTTGCPCTPSPAPNYSPALMRVAAEEVVEVTRVGDFH	2100
2041	NGTMRIIVGPRTCRNMWSGTTFPINAYTTGCPCTPLPAPNYTPALMRVSAEEVEIRRVGDFH	2100
2101	YVTGWTITDNVKCPQVPAPEFFTEVDGVRLLHRVAPACKPLLRDDVTFQVGLNOYLVGSOL	2160
2101	YVTGWTITDNKCPQVPSPEFFTELDGVRLLHRFAPCKPLLRREVEVFRUGLHYHYPVGSOL	2160
2161	PCEPEPDVTLTMSLTDPSHITAETAKRRILARGSPFSLASSASQSLAPSJKATCTTHHD	2220
2161	PCEPEPDVAVLTMSLTDPSHITAEAGRRILARGSPFSPMASSSSQSLAPSJKATCTANHD	2220
2221	SPDADLLEANLLWRQEMGNIITRVESBKNVVIILDSPEPLHAEGDFREISVAABILRKSRK	2280
2221	SPDAELIEANLLWRQEMGNIITRVESBKNVVIILDSFDPLVAEEDEREISVPABILRKSRR	2280
2281	FPSALPTWARPDYNNPPLLESWKPDVYVPVYVHGCPLPPTKAPIPPPRRKRTVVVLTESNV	2340
2281	FAQALPWARPDYNNPPLVETWKKPDVEPPVYVHGCPLPPPSPPVPPPPRKKRTVVLTESTL	2340
2341	SSALAEIATKTFGSSGSSAVDSGTATALPDALSDDGKGDVSESYSGMPPLEGEQDPDL	2400
2341	STALAEIATKSGSSSTSGITGONTTITSSFPAPSGGCPDPSDARSYSMPPLEGEQDPDL	2400

2401 SDGWSWTVSEEA-SEDVVCSSMYTWTGALITPCAAEESKLPINPLNSLLRHHNNVYAT 2459
2401 SDGWSWTVSSEAGTEDVVCSSMSYMTGALVTPCAAEEQKLPINALNSLLRHHNLVYST 2460
2460 TSRSASLRQKKVTFDRLOVLDHYRDLVEMKAKASTVKAKLISIEACKLTPPHSAKSK 2519
2461 TSRSACQRQKKVTFDRLOVLDHYRDLVEMKAKASTVKAKLISIEACKLTPPHSAKSK 2520
2520 FVGAKDVRNLSRAVNHRSVWEDLEDETETIDTITMAKSVFVCQPEKGRKPARLI 2579
2521 FVGAKDVRNLSRAVNHRSVWEDLEDETETIDTITMAKSVFVCQPEKGRKPARLI 2580
2580 VFPDLGVRUCEKMAVDVSTLPQAVMGSSYGFQYSPKQVFEFLVNTWKSCKCPMGFSYD 2639
2581 VFPDLGVRUCEKMAVDVSTLPQAVMGSSYGFQYSPKQVFEFLVNTWKSCKCPMGFSYD 2640
2640 TRCFDSTVTSRDIRVESIYQCCDLAPAEARQAIRSLTERLYIGGPLTNSGQNGYRRCR 2699
2641 TRCFDSTVTSRDIRVESIYQCCDLAPAEARQAIRSLTERLYIGGPLTNSGQNGYRRCR 2700
2700 ASGLVLTSCGNTLTCYLKATAACRAAKLODCTMLVNGDDLWVLCESAGTOEDAAALRAFT 2759
2701 ASGLVLTSCGNTLTCYLKATAACRAAKLODCTMLVNGDDLWVLCESAGTOEDAAALRAFT 2760
2760 EAMTRYSAAPGDPQPPQDYDELELITSCSSNVSVAHADSGKRVYVYLTTRDPTTPLARAAWETA 2819
2761 EAMTRYSAAPGDPQPPQDYDELELITSCSSNVSVAHADSGKRVYVYLTTRDPTTPLARAAWETA 2820
2820 RHTPINSWLGNIIMYAPTLMWMLTHTFFSILLAOBQLEKALDCQIYGACYSIEPLDLP 2879
2821 RHTPINSWLGNIIMYAPTLMWMLTHTFFSILLAOBQLEKALDCQIYGACYSIEPLDLP 2880
2880 QIIRLHGLSAFTLHVSYPGEINRVASCLRLKGLVPPLTRWHRARSVRRAKLLSOGGRAAT 2939
2881 PIIOQLHGLSAFTLHVSYPGEINRVASCLRLKGLVPPLTRWHRARSVRRAKLLSOGGRAAT 2940
2940 CGRYLFNVAWRTKLTPIPAASOLDLSGWFVAGYSGGDIYHSLSRAPRWFPLCLLLLS 2999
2941 CGRYLFNVAWRTKLTPIPAASOLDLSGWFVAGYSGGDIYHSLSRAPRWFPLCLLLLS 3000
3000 VGVGYLLPNR 3010
3001 AGVGYLLPNR 3011

RESULT 13

US-11-126-662-1
; Sequence 1, Application US/11126662
; Publication No. US20050221297A1
; GENERAL INFORMATION:
; APPLICANT: TAKAHAMA, Y.
; SHIRASHI, J.
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS
; C VIRUS INFECTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORBSTER
; STREET: 2000 Pennsylvania Avenue, NW
; City: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/126,662
; FILING DATE: 10-May-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/850,328
FILING DATE: 02-MAY-1997
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mays, Thomas D
REGISTRATION NUMBER: 34,524
REFERENCE/DOCKET NUMBER: 32273-20004.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3011 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-11-126-662-1

Query Match 88.3%; Score 14130.5; DB 6; Length 3011;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 2566; Conservative 225; Mismatches 219; Indels 1; Gaps 1;
Qy Db
1 MSTNPKPQRKTKENTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVNATRKASERSQPRG 60
1 MSTNPKPQKKNENTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVNATRKASERSQPRG 60
61 RRQPIPKARRPEGRNAQPGYPWPLVGNNEGLHAGWLLSPRSPSPSGPTDPRRSNGLG 120
61 RRQPIPKARRPEGRKTWAQPGYPWPLVGNNEGLHAGWLLSPRSPSPSGPTDPRRSNGLG 120
121 KVDTLTTCGADLMGYIPLVGA PLGGAARALAHGVRLVLEGVNATGNLPCSSFSIFLLA 180
121 KVDTLTTCGADLMGYIPLVGA PLGGAARALAHGVRLVLEGVNATGNLPCSSFSIFLLA 180
181 LLSCLTIPASAYEVRNVSGIYHVNTDCSNSIYVEAADVIMHTPGCVPCVQEGNSSRCWV 240
181 LLSCLTIPASAYVRNVSGIYHVNTDCSNSIYVEAADVIMHTPGCVPCVQEGNSSRCWV 240
241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSSAMYVGDLCGSIPLVSQLPFTSPRRHET 300
241 AMPTVATRDGKLPATQLRRHIDLVLGSA TLCSALYVGDLCGSIPLVSQLPFTSPRRHET 300
301 VQDCNCSIYPGHVSGHRMAMWMMNWSPTTALVVSQILRIPOAVVDMVAGAHWGLAGLA 360
301 TQGCNCSIYPGHITGHRMAMWMMNWSPTTALVVSQILRIPOAILDMITAGAHWGLAGIA 360
361 YYSVMVGNWAKVLI VALLFAGVDGETHTTGRVAGHTTSGTSLFSSGASQIKQLVNTNGSW 420
361 YYSVMVGNWAKVLI VALLFAGVDGETHTTGRVAGHTTSGTSLFSSGASQIKQLVNTNGSW 420
421 HINRTALNCNDSLQTFGAALFYAHKFNSSGCCPERMASCRPIDWFAQGWGPITYTKPNSS 480
421 HLNSTALNCNDSLNTGWLGLFYHKKFNSSGCCPERLASCRPLTDFDQGWGPISYANGSGP 480
481 DQRPYCHYAPRCPGVVPASQVCGPVYCFPTSPVVVGTTRDSGVPTTYSWGNEDTVMJLN 540
481 DQRPYCHYPPKPGCI VPAKSVCGPVYCFPTSPVVVGTTRDSGVPTTYSWGNEDTVMJLN 540
541 NTRPPQGNWFCCTWNSTGFTKTCGGPPCNTGGVGNRTLICPTDFCRKHGPATYTKCSG 600
541 NTRPPQGNWFCCTWNSTGFTKTCGGPPCNTGGVGNRTLICPTDFCRKHGPATYTKCSG 600
601 PWTLPRLVDYPRYLWHYPCTLNFSIFKVRMYVGVGVEHRLAACNWTGERCNLDRDRS 660
601 PWTLPRLVDYPRYLWHYPCTLNFSIFKVRMYVGVGVEHRLAACNWTGERCNLDRDRS 660
661 ELSPLLLSTTEWQILPCAFATTLPALSTGLIHLHQNIVDVQVLYGVGSFAVSPFAIKWEYL 720
661 ELSPLLLSTTEWQILPCAFATTLPALSTGLIHLHQNIVDVQVLYGVGSFAVSPFAIKWEYL 720

Qy	721	LLFLLLDARVCA	CLWMLLIAQBAE	ALENLVLNAA	SVAGAHGILS	FLVFFCAAWYIK	780												
Db	721	LLFLLLDARVCS	CLWMLLI	ISQAEAALE	NLVI	LNAASLAGTHGLSV	FLVFFCFAWYIKG 780												
Qy	781	RLAPGAAVAF	GVWPLLLLLL	ALLPPRAYAL	DREMAASCGG	ANVLGLVFLTL	SPYKVFLT 840												
Db	781	KWPGAVYTF	GVWPLLLLLL	ALLPQRAYAL	DTVEAASCGG	VVLVGLMALTL	SPYKRYIS 840												
Qy	841	RLIWLQVFI	TRAEAHQVW	VPPLNVRGGR	DAIILLTCAV	HPELIFOITKLL	LAAILPLM 900												
Db	841	WCLWNLQV	FLTRVZAO	LHVMIPLNVRGGR	DAVILLMCAV	HTLVDFDITKLL	LAVFGPLW 900												
Qy	901	VLQAGITRV	PVYFVBAQ	GLIRACMLVR	KVAGHYVQMV	FMKLGALTGT	YVYNNHLTLR DWA 960												
Db	901	ILQASLLK	VPYFVRVQ	GLRFLCAR	KWIGHYVQW	VIKLGALTGT	YVYNNHLTLR DWA 960												
Qy	961	HAGLRDLA	VAVEPVPV	FSAMETKVIT	WGADTAAC	GDIIILGIP	SVSARRKEIFLGPADSLG 1020												
Db	961	HNGLRDLA	VAVEPVPV	FSQMETKLT	TWGADTAAC	GDIIINGL	IPVSARRGREILLPADGMVS 1020												
Qy	1021	QGWELLAP	ITAYSOOT	RGVLGCIT	STLUTGRD	KNOVEGEV	QVSVSTATOSFIATCINGVCWT 1080												
Db	1021	KGWELLAP	ITAYAOQT	RGLGCIIT	SLUTGRD	KNOVEGEV	QIVSTAATFIATCINGVCWT 1080												
Qy	1081	VYHAGSKT	LAPKPGPIT	QMTYNTVD	LVGWQAP	PGARSMT	PCSCGSSDLYLVTRHADVI 1140												
Db	1081	VYHAGTET	IASPKGPV	QMTYNTVD	QLVGWPA	PQGSRLT	TPCTCGSSDLYLVTRHADVI 1140												
Qy	1141	PVRRGDSR	GLLSPR	VSYLKSGS	GPLCP	SGHVHV	GFRAAAVCTRGVAKAVDFIPVES 1200												
Db	1141	PVRRGDSR	GLLSPR	PISYLKSGS	GPLCP	PAGHAV	IGIFRAAAVCTRGVAKAVDFIPVEN 1200												
Qy	1201	METMRS	SPVFDNST	PPAVPOT	FOVHLHAPT	SGSKTKV	PAAVAAQYKVLVLNPSVAA 1260												
Db	1201	LETMRS	SPVFDNSP	PPVPOS	FOVHLHAPT	SGSKTKV	PAAVAAQYKVLVLNPSVAA 1260												
Qy	1261	TLFGAYMS	KAHGID	PNIRTV	RTITGGST	ITYGKFL	ADGCGSGAYDIIICDBCHS 1320												
Db	1261	TLFGAYMS	KAHGID	PNIRTV	RTITGSP	ITYGKFL	ADGCGSGAYDIIICDBCHS 1320												
Qy	1321	TDSTTILG	IGFVLDOA	ETAGAR	LVLATAT	PPGSVTV	PHPNIEBIGLSNNGEIPFYKAI 1380												
Db	1321	TDATSLIG	ITVLDOA	ETAGAR	LVLATAT	PPGSVTV	PHPNIEIEVALSTTGEIPFYKAI 1380												
Qy	1381	PIEAIKGR	HLIFCHSK	KKDELA	AKLTGL	GLNAVAY	YRGLDVSVIIPIDGVVVVATDAL 1440												
Db	1381	PLEVIKGR	HLIFCHSK	KKDELA	AKLVAL	GINAVAY	YRGLDVSVIPTSGDVVVVATDAL 1440												
Qy	1441	MTGTG	DFDSVID	CNTCVT	QVDFSL	DPFTT	ETITTV	PQDAVSRBSRRGRTGRSGIYR 1500											
Db	1441	MTGTG	DFDSVID	CNTCVT	QVDFSL	DPFTT	ETITTV	PQDAVSR	TQRGRTRGKPGIYR 1500										
Qy	1501	FVTPG	ERP	SGMFDSS	VLCECY	DACAGW	VELTPA	ETSVRL	RAYLNT	PGLPVCQD	HEFWS 1560								
Db	1501	FVAPG	ERP	SGMFDSS	VLCECY	DACAGW	VELTPA	ETTVRL	RAYNNT	PGLPVCQD	HEFWEG 1560								
Qy	1561	VFTGLTH	IDAHLF	SOTKQAGN	FPYLVA	QATVCARA	QAPPPS	WDQW	MKCLIRL	KPTLHG 1620									
Db	1561	VFTGLTH	IDAHLF	SOTKQAGN	FPYLVA	QATVCARA	QAPPPS	WDQW	MKCLIRL	KPTLHG 1620									
Qy	1621	PTPLL	YRLGAV	ONEVIL	THPTIK	YIMACMS	ADLE	EVWTS	TVVL	UGGV	LAALAA	CYLTG	TSV 1680						
Db	1621	PTPLL	YRLGAV	ONEBIT	THPTVK	YIMT	CMSAD	LE	VWTS	TVVL	UGGV	LAALAA	YCYCLSTGCV 1680						
Qy	1681	VIVGRI	ILSGKPA	VVP	REVLY	QBFDEW	ECCASOL	PYIE	QWOLAE	OFKOKAL	GLLLOTAT 1740								
Db	1681	VIVG	RVLSG	KPAI	I	P	REVLY	RFDEW	ECCS	QHL	PYIE	QWMLAE	OFKOKAL	GLLLOTAS 1740					
Qy	1741	KQAE	AAAP	VVES	KWRAL	ET	FWAK	HMW	NFI	ISGI	OY	LAG	STL	PGN	PAI	AS	MAFT	ASITSP 1800	
Db	1741	ROAE	VI	AP	AVOT	NQ	KLET	FWAK	HMW	NFI	ISGI	OY	LAG	STL	PGN	PAI	AS	MAFT	AAVTSPT 1800

Qy	1801	LTTQNTLLFNILCGWAAQIAPPASAFAVAGAGIAGAVSGIGLKVLDILAGYGAGVA	1861
Db	1801	LTTQNTLLFNILCGWAAQIAPPASAFAVAGAGIAGAVSGIGLKVLDILAGYGAGVA	1860
Qy	1861	GALVAFKVMGSEVPSTEDILVNLPAILSPGALVGVVCAAILRRHHVPGBGAVQVMNRLI	1920
Db	1861	GALVAFKVMGSEVPSTEDILVNLPAILSPGALVGVVCAAILRRHHVPGBGAVQVMNRLI	1920
Qy	1921	AFASRGNHVSPTHYVPESDAAARVTQILSSLTITQLLKRILHOMINEDCSTPCSGSWLRDV	1980
Db	1921	AFASRGNHVSPTHYVPESDAAARVTQILSSLTITQLLKRILHOMINEDCSTPCSGSWLRDI	1980
Qy	1981	WDMICVILSDFTKWLAKMLPQLPGIPFVSCQSGYKGVWRGDGIMHTRCHGCAEITGHVK	2040
Db	1981	WDMICVILSDFTKWLAKMLPQLPGIPFVSCQSGYKGVWRGDGIMHTRCHGCAEITGHVK	2040
Qy	2041	NGSMRIVGPRCTCSNTHGHTPPINAYTGTCTPSPAPNYSPALWRVAEEVVEVTRVCGDFH	2100
Db	2041	NGSMRIVGPRCTCSNTHGHTPPINAYTGTCTPSPAPNYSPALWRVAEEVVEVTRVCGDFH	2100
Qy	2101	YVTGMTTDNVKPCQVPAPPEFFVEVDGVRLLHRVAPACKPLLRDVTFOVGLNOYLVGSOL	2160
Db	2101	YVTGMTTDNVKPCQVPAPPEFFVEVDGVRLLHRVAPACKPLLRDVTFOVGLNOYLVGSOL	2160
Qy	2161	PCPEPDVTVLTSMLTDPDSHITAETAKRRLARGSPPLSSASSASQLSAPLSKATCTTHHD	2220
Db	2161	PCPEPDVTVLTSMLTDPDSHITAETAKRRLARGSPPLSSASSASQLSAPLSKATCTANHD	2220
Qy	2221	SPADLIEANLLWRQEMGNIITRVESENKVLIIDSPFLHAEODEREISVAABILRKSRK	2280
Db	2221	SPADLIEANLLWRQEMGNIITRVESENKVLIIDSPFLHAEODEREISVPABILRKSR	2280
Qy	2281	FPSALPIWARPDYNNPPLLESWKDPDVPVPPVWHGCPPLPTKAPPIPPRRKRTTVLTVESNV	2340
Db	2281	FAQALPVWARPDYNNPPLVETWKPDVPEPPVWHGCPPLPPKSPVPPPRKRTTVLTVESTL	2340
Qy	2341	SSALAEIATKTFSGSSSAVDSGTATALPDASDDGDKGSDVESYSGMPPLEGEQDPDL	2400
Db	2341	STALAEIATKTFSGSSSTSGITGDNWTTTSBEPAPSGPPDSDAESYSGMPPLEGEQDPDL	2400
Qy	2401	SDGSWSTVSEEA-SEDVVCSCMSYWTGALITCAAEESKLPINPLNSNLLRHHNMVYAT	2459
Db	2401	SDGSWSTVSEANAEDVVCSCMSYSMTGALVTPCAABEQKLPINALNSNLLRHHNLVYST	2460
Qy	2460	TSSASILQKQKVTFDRLQVLDDHVRDVLKEMKAKASTVKAKLISIEACKLTPPHSAKSX	2519
Db	2461	TSSACORQKQKVTFDRLQVLDDSHYQDVLKEVKAASKVKANLISVEEACSLTTPPHSAKSX	2520
Qy	2520	FGYGAQDVRNLSRRVNHRSVWEDLLEDTPETPIDTTIMAKSVFVCQPEKGGKRPALRI	2579
Db	2521	FGYGAQDVRCHARKAVTHINSVWKDLLEDNVPTIDTTIMAKNEVFCVQPEKGGKRPALRI	2580
Qy	2580	VFPDLGVRVCEKMALYDVVSTLPOAVMGSSYGFQSPQKQVFEFLVNTWKSKCPMGFSYD	2639
Db	2581	VFPDLGVRVCEKMALYDVVTKLPLAVMGSSYGFQSPGQVFEFLVQAWKSCKTTPMGFSYD	2640
Qy	2640	TRCFDSTVSTESDIRVEESIIYOCDDLAPEARQAIRSLTERLYIGGLPTNSKGONCVYRRCR	2699
Db	2641	TRCFDSTVSTESDIRTEBAIIYOCDDLPOARVALIKSLTERLYVGGPLTNSRGENCVYRRCR	2700
Qy	2700	ASGVLTSCGNTLTICYLKATAACRAAKLODQTMVNGDDLWVTCESAGTOQEDAAALRAFT	2759
Db	2701	ASGVLTSCGNTLTICYLKARAACRAGLODQTMVNGDDLWVTCESAGVOEDAAALRAFT	2760
Qy	2760	EAMTRYSAAPGDPPOPEYDLELITSCSSNVSVAHADSGKRVYVYLTRDPTTPLARAAWETA	2819
Db	2761	EAMTRYSAAPGDPPOPEYDLELITSCSSNVSVAHADGAGKRVYVYLTRDPTTPLARAAWETA	2820
Qy	2820	RHPTPNSWLGNIIMYAPTLWARMILMTHFPSSILLAQOLEKALDCQIYGACYSIEPLDLP	2879
Db	2821	RHPTPNSWLGNIIMFAPTLWARMILMTHFPFSVLJARQOLEQALDCIYGACYSIEPLDLP	2880
Qy	2880	QIETRLHGLSAFTLHYSYSGPEINRVASCLFKLGVPLPRTWRHRARSVRAKLLSOGGBAAT	2939

1621 PTPLLYRLGAVONEVILTHPIKTYIMACMSADLEVVVSTWLVGVGLAALAAAYCLTTGVS 1680
1621 PTPLLYRLGAVONEVILTHPIKTYIMTMSADLEVVVSTWLVGVGLAALAAAYCLTGCV 1680
1681 VIVGRIILSGKPAVVPDREVLVYQEFDEMEECASQIPYIEQGMOLAEQFKQKALGLLQAT 1740
1681 VIVGRIILSGKPAIIPDREVLVYQEFDEMEECASQIPYIEQGMOLAEQFKQKALGLLQAT 1740
1741 KOAEAAAAPVESKWALETFWAKHWNFTSGIYLAGLSTLPCNPALIASLMAFTASITSP 1800
1741 RHAEVITPAVQTNWQKLEFWAKHWNFTSGIYLAGLSTLPCNPALIASLMAFTAAVTSP 1800
1801 LTTQNTLPLNIIIGGWAAQAPPSAASAFVAGIAGAAVSGIGLKVLDVILLAGYAGVA 1860
1801 LTTGQTLPLNIIIGGWAAQAPGAATAFVAGIAGAAVSGIGLKVLDVILLAGYAGVA 1860
1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVOMNRLI 1920
1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVOMNRLI 1920
1921 AFASRGNHVSPTHYPVSDAAARVTOILLSSLTITQLLKHQWINEDECSTPCSGSLRDL 1980
1921 AFASRGNHVSPTHYPVSDAAARVTAIILSSLTITQLLRLHQWISSECTTPCSCSLRDI 1980
1981 WDWICTVLTFDTWLTQSKLLPRLPGVPLSCORGKGVWRGDMQTTCPGCAQIAGHVK 2040
1981 WDWICEVLSDFKTKAKLMPQPGIPFVSCORGKGVWRGDMQTTCPGCAQIAGHVK 2040
2041 NGSMRIVGPRTCNTWHGTFPINAITYTGTCTSPAPNTSRALMRVAAEYEVETRVGDFH 2100
2041 NGTMRIVGPRTCNTWHGTFPINAITYTGTCTPLPAPNYKFAKLRVSAEYEVETRVGDFH 2100
2101 YVTGMTDNNVCKQVAPAPPTVEVDGVRHLRYAPACKPELLREDYTFQVGLNQYLVGSOL 2160
2101 YVSGMTDNLKCPQITSPSEFFTELDGVRHLRFAPPKPELLREEVSVFVGLHEYVPGSOL 2160
2161 PCEPPEPDVTLTSMLTDRSHITAETAKRLRAGSPSPSLASSASQLSAPSLKATCTTHD 2220
2161 PCEPPEPDVTLTSMLTDRSHITAETAKRLRAGSPSPSLASSASQLSAPSLKATCTAHND 2220
2221 SPDADLEIANLWRQBMGNITRVSEKNVILDSPEPLHAEGDREISVAAEILLRKSK 2280
2221 SPDADLEIANLWRQBMGNITRVSEKNVILDSFDPLVAEEDEREYSVPAEILLRKSK 2280
2281 FPSALPIWARPDPYNPLLESWKDDPVVPPVHGCPLPPTKAPPIPPPRKRTVVLTESV 2340
2281 FARALPVMARPDPYNPLVETWKKKDPYPPVHGCPLPPTKAPPIPPPRKRTVVLTESTL 2340
2341 SSALAEATKTFGSSGSSAVDSGTATALPDASDDGDKGSDVESYSSMPLEGEPCDDPL 2400
2341 STALAEATKTFGSSGSSGSDGTATALPDASDDGDKGSDVESYSSMPLEGEPCDDPL 2400
2401 SDGSMSTVSEEA- SEDVVCSSSYWTGALITPCAAEESKLPINPLNSLLRHHNNVAT 2459
2401 SDGSMSTVSGADTEDVVCSSSYWTGALVTPCAAEQKLPINALNSLLRHHNLVYST 2460
2460 TSRSASLRQKVTDFRLQVLDHYRDVLEKMAKASTVAKLLSTEEACKLTPPHSAKSK 2519
2461 TSRSACQKQKVTDFRLQVLDHYQVLEKMAKASTVAKLLSTEEACKLTPPHSAKSK 2520
2520 FGVGAKDVNLSRAVNHITRSVWEDLLETEPTIDTTINAKSEVFCVQPEKGRKPARLI 2579
2521 FGVGAKDVNLSRAVNHITRSVWEDLLETEPTIDTTINAKSEVFCVQPEKGRKPARLI 2580
2580 VFPDLGVRVCEKXALYDVVSTLTPQAVGSSYGFQSPKORVFLVNTWKSKECPMGFSYD 2639
2581 VFPDLGVRVCEKXALYDVVSTLTPQAVGSSYGFQSPKORVFLVNTWKSKECPMGFSYD 2640
2640 TRCFDSTVTSIEDIRVEESYQCCDLAPARQAIRSLTERLYTGGPLTNSKQNCQYRROR 2699
2641 TRCFDSTVTSIEDIRVEESYQCCDLAPARQAIRSLTERLYTGGPLTNSKQNCQYRROR 2700

2700 ASGVLTTSCGNTLTCYLKATAACRAAKIQDCTMLVNGDDLVIICESAGTQEDAAALRAFT 2759
2701 ASGVLTTSCGNTLTCYLKATAACRAAKIQDCTMLVNGDDLVIICESAGVQEDAAALRAFT 2760
2760 EAMTRYISAPPDPPQPEYDLELITSCSSNVSAVDHAGSKRVYLYLTRDPTTPLARAWEA 2819
2761 EAMTRYISAPPDPPQPEYDLELITSCSSNVSAVDHAGSKRVYLYLTRDPTTPLARAWEA 2820
2820 RHTPINSWLGNIINVAPTLWARMILMTHFFSILLAOEOLKALDCQIYGACYSISPLDLP 2879
2821 RHTPINSWLGNIINVAPTLWARMILMTHFFSVLTARDQEQALNCEIYGACYSISPLDLP 2880
2880 QIIRLHGLSFTLHSHSPGSEINRVASCLRLKGVPPPLFTRHRRARSVRAKLLSQCGRAAT 2939
2881 PIQRLHGLSFTLHSHSPGSEINRVASCLRLKGVPPPLFTRHRRARSVRARULLSRGAAI 2940
2940 CGRYLFNVAWTKLKLTPIPAASQDLDSGWFVAGYSGGDIYHSLSRARPRPPLCLLLLS 2999
2941 CGRYLFNVAWTKLKLTPIPAASQDLDSGWFVAGYSGGDIYHSLSRARPRPPLCLLLLS 2999
3000 VGVGIYLLPNR 3010
3001 AGVGIYLLPNR 3011

RESULT 15
US-09-747-419-20
; Sequence 20, Application US/09747419
; Patent No. US20020155582A1
; GENERAL INFORMATION:
; APPLICANT: Yi, MinKyoung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007.0101
; CURRENT APPLICATION NUMBER: US/09/747,419
; CURRENT FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 20
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polyprotein
US-09-747-419-20

Query Match 88.2%; Score 14124.5; DB 3; Length 3011;
Beat Local Similarity 85.5%; Pred. No. 0;
Matches 2573; Conservative 218; Mismatches 219; Indels 1; Gaps 1;
QY 1 MSTNPKPQRTKRNTRRRPQDKVPGGQIVGGVYLLPRRGLGVRAATRKASERSQPRG 60
DB 1 MSTNPKPQRTKRNTRRRPQDKVPGGQIVGGVYLLPRRGLGVRAATRKASERSQPRG 60
QY 61 RROPIPKARRPGRWAQGYWPVLYGNEGKWAAGLLSPRGSRSWGPDPDRRSRNLG 120
DB 61 RROPIPKARRPGRWAQGYWPVLYGNEGKWAAGLLSPRGSRSWGPDPDRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVAPLGGAAALAHGVRLVEDGVNATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVAPLGGAAALAHGVRLVEDGVNATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIPASAYEVNRVSGIYHVTNDCNSSIYVEAADVIMHTPGCVPCVQEGNSRCWV 240
DB 181 LLSCLTIPASAYEVNRVSGIYHVTNDCNSSIYVEAADVIMHTPGCVPCVQEGNSRCWV 240
QY 241 ALTPTLAARNASVPTTTRRRHVDLLVGTAAFCSAMVGDLCGSIPLVSOLFSPRRHET 300
DB 241 ALTPTLAARNASVPTTTRRRHVDLLVGTAAFCSAMVGDLCGSIPLVSOLFSPRRHET 300
QY 301 VQDCNCSIYPGHVSGHRRMAWDMNMNWSPTTALVVSQLLRIPOAVVDMVAGAHGWGLAGLA 360

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 08:32:58 ; Search time 13 Seconds
(without alignments)
1108.683 Million cell updates/sec

Title: US-09-662-454-3
Perfect score: 16009
Sequence: 1 MSTNPKPQRTKNTNRPO.....FPLCLLLSVGVGYLLPNR 3010

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA New:*
1: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pdb:*
2: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pdb:*
3: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pdb:*
4: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pdb:*
5: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pdb:*
6: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pdb:*
7: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pdb:*
8: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	11539	72.1	2280	7	US-11-022-562-211
2	10106	63.1	1985	7	US-11-119-330-2
3	2979	18.6	637	7	US-11-154-324-4
4	1943.5	12.1	446	7	US-11-117-667-2
5	218	1.4	3433	6	US-10-714-781A-67
6	171	1.1	1043	6	US-10-821-234-1055
7	142	0.9	771	6	US-10-949-720-389
8	131	0.8	401	6	US-10-949-720-419
9	130	0.8	5405	7	US-11-108-172-1116
10	129	0.8	4384	6	US-10-821-234-1120
11	127	0.8	522	6	US-10-949-720-425
12	127	0.8	537	6	US-10-949-720-424
13	127	0.8	555	6	US-10-949-720-387
14	127	0.8	570	6	US-10-949-720-386
15	127	0.8	570	6	US-10-949-720-412
16	127	0.8	987	6	US-10-949-720-395
17	127	0.8	990	6	US-10-821-234-1201
18	124	0.8	20	6	US-10-962-145C-8
19	122	0.8	20	6	US-10-962-145C-16
20	120	0.7	1493	7	US-11-004-057-21
21	119	0.7	488	6	US-10-467-657-2362
22	119	0.7	7968	7	US-11-186-731-5
23	118	0.7	901	7	US-11-082-389-430
24	117.5	0.7	795	6	US-10-821-234-1675
25	116	0.7	20	6	US-10-962-145C-11

26	116	0.7	20	6	US-10-962-145C-17	Sequence 17, Appli
27	115.5	0.7	747	7	US-11-018-018-1	Sequence 1, Appli
28	115.5	0.7	747	7	US-11-047-757-1	Sequence 1, Appli
29	115	0.7	898	7	US-11-174-150-43	Sequence 43, Appli
30	115	0.7	1311	6	US-10-509-422-5	Sequence 5, Appli
31	113.5	0.7	984	7	US-11-055-822-508	Sequence 508, App
32	113.5	0.7	984	7	US-11-055-822-594	Sequence 594, App
33	113.5	0.7	1267	7	US-11-109-156-35	Sequence 35, Appli
34	113	0.7	4419	6	US-10-821-234-1155	Sequence 1155, Ap
35	111	0.7	20	6	US-10-962-145C-7	Sequence 7, Appli
36	111	0.7	1637	6	US-10-821-234-1204	Sequence 1204, Ap
37	110.5	0.7	826	7	US-11-055-822-214	Sequence 214, App
38	110.5	0.7	826	7	US-11-055-822-712	Sequence 712, App
39	110.5	0.7	833	7	US-11-055-822-212	Sequence 212, App
40	110.5	0.7	833	7	US-11-055-822-710	Sequence 710, App
41	110.5	0.7	1221	6	US-10-858-730-222	Sequence 222, App
42	109.5	0.7	495	7	US-11-074-176-266	Sequence 266, App
43	109.5	0.7	667	6	US-10-821-234-1477	Sequence 1477, Ap
44	109.5	0.7	748	6	US-10-821-234-888	Sequence 888, App
45	109.5	0.7	957	7	US-11-108-172-1065	Sequence 1065, Ap

ALIGNMENTS

RESULT 1
US-11-022-562-211
; Sequence 211, Application US/11022562
; Publication No. US20050249742A1
; GENERAL INFORMATION:
; APPLICANT: Ruprecht, Ruth M.
; APPLICANT: Shisong, Jiang
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; TITLE OF INVENTION: A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
; FILE REFERENCE: DFN-043CN
; CURRENT APPLICATION NUMBER: US/11/022,562
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: PCT/US03/20322
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/3927718
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 211
; LENGTH: 2280
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-11-022-562-211

Query Match	72.1%	Score	11539;	DB	7;	Length	2280;
Best Local Similarity	93.7%	Pred. No.	0;				
Matches	2137;	Conservative	69;	Mismatches	74;	Indels	0;
Gaps	0;						
Qy	1	MSTNPKPQRTKNTNRPODVKPPGGQIVGVYLLPRGPRIGVTRKASERSQPRG	60				
Db	1	MSTNPKPQRTKNTYRPPQDVKPPGGQIVGVYLLPRGPRIGVTRKASERSQPRG	60				
Qy	61	RRQIPKARPEGRWAQPGYPWPLYGNEGIGWAGLLSPRGSRPMSGPTDPRRSNLG	120				
Db	61	RRQIPKARPEGRWAQPGYPWPLYGNEGIGWAGLLSPRGSRPMSGPTDPRRSNLG	120				
Qy	121	KVIDTLTCGFDLGMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPCGSFIFLLA	180				
Db	121	KVIDTLTCGFDLGMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPCGSFIFLLA	180				
Qy	181	LLSCLTTPASAYEVRNYSYIYHVNDSCNSIVVEADVINHTPCVCVCGESSRCWV	240				
Db	181	LLSCLTTPASAYVRNYSYIYHVNDSCNSIVVEADVINHTPCVCVCGESSRCWV	240				
Qy	241	ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSSAMYVGLCGSIFLVLSQFTFSPRRHET	300				
Db	241	ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSSAMYVGLCGSIFLVLSQFTFSPRRHET	300				

QY 301 VQDCNCSYPGHVSGRHMAWMMNWSPTTALVVSQLLRIPQAVVDMVAGHWGLAGLA 360
DB 301 VQDCNCSYPGHVSGRHMAWMMNWSPTTALVVSQLLRIPQAVVDMVAGHWGLAGLA 360
QY 361 YYSVMGNKAKULIALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQKQLVNTNGSW 420
DB 361 YYSVMGNKAKULIALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQKQLVNTNGSW 420
QY 421 HNRRTALNCNDSLOTFQFAALFYAHKENSNGCPCPERMASCRIPTDFAQCGMPITYTKPNSS 480
DB 421 HNRRTALNCNDSLOTFQFAALFYAHKENSNGCPCPERMASCRIPTDFAQCGMPITYTKPNSS 480
QY 481 DORPYCWHYAPRCGVPYASQVCGPVYCFPTSPVVVGTTRDSGVPTYSWGENETDVMMLN 540
DB 481 DORPYCWHYAPRCGVPYASQVCGPVYCFPTSPVVVGTTRDSGVPTYSWGENETDVMMLN 540
QY 541 NTRPPQGNWFGCTWNNSTGFTKTCGPPCNIIGVGNLTLCTDCFRKHPEATYTKCGSG 600
DB 541 NTRPPQGNWFGCTWNNSTGFTKTCGPPCNIIGVGNLTLCTDCFRKHPEATYTKCGSG 600
QY 601 PMLTRCLVDYBYRLWHYPCNTLNFSEFKVMYVGVGVEHNLNAACNWTGERCNLEDRDRS 660
DB 601 PMLTRCLVDYBYRLWHYPCNTLNFSEFKVMYVGVGVEHNLNAACNWTGERCNLEDRDRS 660
QY 661 ELSPLLSTTEWQIILPCAFPTLIPALSTGLIHLHQNIVDVQYLYGVGSAPVSPAIAKWEYIL 720
DB 661 ELSPLLSTTEWQIILPCAFPTLIPALSTGLIHLHQNIVDVQYLYGVGSAPVSPAIAKWEYIL 720
QY 721 LFLFLLLADARVACACILWMLLIIAQAEAALENVLVNAASVAGAHGILSLFLVFFCAAWYIKG 780
DB 721 LFLFLLLADARVACACILWMLLIIAQAEAALENVLVNAASVAGAHGILSLFLVFFCAAWYIKG 780
QY 781 RLARGAAYFYGVWPLILILILALPRAYALDREMAASCGGAVLGLVFLTSLSPYKVELT 840
DB 781 RLARGAAYFYGVWPLILILILALPRAYALDREMAASCGGAVLGLVFLTSLSPYKVELT 840
QY 841 RLILWLOVYFTRAEAHMQVWVPLNVGRGGRDAIILLTCAVHPPELIDITKJLLAILGLPLM 900
DB 841 RLILWLOVYFTRAEAHMQVWVPLNVGRGGRDAIILLTCAVHPPELIDITKJLLAILGLPLM 900
QY 901 VLQAGITRVPYFVRAQGLIRACMLVRKVAGHYVQVMFKLGAALTGYTYVYNHLTPLRDWA 960
DB 901 VLQAGITRVPYFVRAQGLIRACMLVRKVAGHYVQVMFKLGAALTGYTYVYNHLTPLRDWA 960
QY 961 HAGLRDLAVAVPVPFVFSANETKVIITWGADTAACGDIILGLPVSARRGKEIFIGPADSLBG 1020
DB 961 HAGLRDLAVAVPVPFVFSANETKVIITWGADTAACGDIILGLPVSARRGKEIFIGPADSLBG 1020
QY 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVSTATQSFLATCINGVCWT 1080
DB 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVSTATQSFLATCINGVCWT 1080
QY 1081 VYHGAGSKTLAGPKGITQMYTNVDLDLVGWAQPPGARSMTPCSGSSDLXLYVTRHADVI 1140
DB 1081 VYHGAGSKTLAGPKGITQMYTNVDLDLVGWAQPPGARSMTPCSGSSDLXLYVTRHADVI 1140
QY 1141 PYRRGDSRGSLLSPRPVSYLKGSSGCPILCPSHVGVGFRAVCTRGYAKAVDFIPVES 1200
DB 1141 PYRRGDSRGSLLSPRPVSYLKGSSGCPILCPSHVGVGFRAVCTRGYAKAVDFIPVES 1200
QY 1201 METTMRSPVFTDNSTPPAVPQTFQVAHLHAPGTSGSKSTKVPAAAYAQQYKVLVLPNSVAA 1260
DB 1201 METTMRSPVFTDNSTPPAVPQTFQVAHLHAPGTSGSKSTKVPAAAYAQQYKVLVLPNSVAA 1260
QY 1261 TLGFGAYMSKAHGIDPNIRGTVRTITTGSIITYSTYKFLADGGCSGGAYDIIICDECHS 1320
DB 1261 TLGFGAYMSKAHGIDPNIRGTVRTITTGSIITYSTYKFLADGGCSGGAYDIIICDECHS 1320
QY 1321 TDSITLIGTGLVDQAEAGARLVLATATPPGSVTVPHNTEIEIGLSNNGEIPFYGKAI 1380
DB 1321 TDSITLIGTGLVDQAEAGARLVLATATPPGSVTVPHNTEIEIGLSNNGEIPFYGKAI 1380
QY 1381 PIEAKGRHILFCHSKKKDELAALKLGLGLNAVAYYRGLDVSVPPIGDVVVVVATDAL 1440

DB 1381 PIEAKGRHILFCHSKKKDELAALKLGLGLNAVAYYRGLDVSVPPIGDVVVVVATDAL 1440
QY 1441 MTGFTGDPDSVIDNCNTCVTOTVDFSLDPTFTIETTTVPQDAVRSORSORGRTRGRSGIYR 1500
DB 1441 MTGFTGDPDSVIDNCNTCVTOTVDFSLDPTFTIETTTVPQDAVRSORSORGRTRGRSGIYR 1500
QY 1501 FVTGPERSGMFDSSVLCCEYDAGCAWYELTAPETSVRRLAYLNTFPLPVCDHLEFWES 1560
DB 1501 FVTGPERSGMFDSSVLCCEYDAGCAWYELTAPETSVRRLAYLNTFPLPVCDHLEFWES 1560
QY 1561 VFTGLTHIDAHFLSOTKOAGDNFPYLVAIQATVCARAQAPPPSWDMKKLIRLKPRTLHG 1620
DB 1561 VFTGLTHIDAHFLSOTKOAGDNFPYLVAIQATVCARAQAPPPSWDMKKLIRLKPRTLHG 1620
QY 1621 PTPLLYRLGAVONEVILTHPIITKYMACHSADLEVVSTVWLVGVLAAALAYCLTTGVS 1680
DB 1621 PTPLLYRLGAVONEVILTHPIITKYMACHSADLEVVSTVWLVGVLAAALAYCLTTGVS 1680
QY 1681 VIVGRIILSGKPAVVPDREVLYOEFDEMEECASQLPYIEQGMQLAEQFKQKALGLLOTAT 1740
DB 1681 VIVGRIILSGKPAVVPDREVLYOEFDEMEECASQLPYIEQGMQLAEQFKQKALGLLOTAT 1740
QY 1741 KOAEAAAPVVESKRALETFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
DB 1741 KOAEAAAPVVESKRALETFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
QY 1801 LTTQNTLLFNILGHWAAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDIILAGYGAGVA 1860
DB 1801 LTTQNTLLFNILGHWAAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDIILAGYGAGVA 1860
QY 1861 GALVAFKVMSEVPSTEDLVNLLPAILSPGALVGVVCAAIILRRHVGPGEAVQWNNRLI 1920
DB 1861 GALVAFKVMSEVPSTEDLVNLLPAILSPGALVGVVCAAIILRRHVGPGEAVQWNNRLI 1920
QY 1921 AFASRGNHVSPTHYVPESDAAARVTQIILSSLTITQLKRLHQMINEDECSPTPCSGMWLKV 1980
DB 1921 AFASRGNHVSPTHYVPESDAAARVTQIILSSLTITQLKRLHQMINEDECSPTPCSGMWLKV 1980
QY 1981 WDWICTVLTDKRWLQSLKLPRLPGVPFLSCORGKGVWVRGDMQOTCCGCAQIAGHVK 2040
DB 1981 WDWICTVLTDKRWLQSLKLPRLPGVPFLSCORGKGVWVRGDMQOTCCGCAQIAGHVK 2040
QY 2041 NGSMRIVGPRTCSNTHWGTFFPINAYTGPCTSPAPNSRALLRWVAAEEYVEVTRVGDH 2100
DB 2041 NGSMRIVGPRTCSNTHWGTFFPINAYTGPCTSPAPNSRALLRWVAAEEYVEVTRVGDH 2100
QY 2101 VYTGMTTNDVKPCQVPAPEPFEVDGVRHLHRYAPACKPLLEDVTFQVGLNQYLVGSOL 2160
DB 2101 VYTGMTTNDVKPCQVPAPEPFEVDGVRHLHRYAPACKPLLEDVTFQVGLNQYLVGSOL 2160
QY 2161 PCEPEPDVTLTSMLTDRSHITAEAKRLARGSPPSLASSASQISAPSLKATCTTHD 2220
DB 2161 PCEPEPDVTLTSMLTDRSHITAEAKRLARGSPPSLASSASQISAPSLKATCTTHD 2220
QY 2221 SPDADLIEANLLWROEMGNITRVESENKVILDSFEPLHAEGDEREISVAAEILKRSK 2280
DB 2221 SPDADLIEANLLWROEMGNITRVESENKVILDSFEPLHAEGDEREISVAAEILKRSK 2280

RESULT 2

US-11-119-330-2
; Sequence 2, Application US/11119330
; Publication No. US20050260568A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Min
; APPLICANT: Lemm, Julie A.
; APPLICANT: O'Boyle, Donald R.
; APPLICANT: Nowar, Peter
; TITLE OF INVENTION: HEPATITIS C VIRUS ASSAYS
; FILE REFERENCE: 10283 NP
; CURRENT APPLICATION NUMBER: US/11/119,330
; CURRENT FILING DATE: 2005-04-29

;; PRIOR APPLICATION NUMBER: 60/567,270
;; PRIOR FILING DATE: 2004-04-30
;; PRIOR APPLICATION NUMBER: 60/568,590
;; PRIOR FILING DATE: 2004-05-06
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: Patent in version 3.3
;; SEQ ID NO 2
;; LENGTH: 1985
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: HCV Replicon
US-11-119-330-2

Query Match 63.1%; Score 10106; DB 7; Length 1985;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 1903; Conservative 41; Mismatches 41; Indels 0; Gaps 0;

QY	1026	LAPITAYSQOTRGVLGCIITSLTGRDNQVEGEVQVVSSTATQSLFATCINGVCWTVYHGA	1085
DB	1	MAPITAYSQOTRGLGCIITSLTGRDENQVEGEVQVVSSTATQSLFATCINGVCWTVYHGA	60
QY	1086	GSXTLAGPKGIPTQMTYNVDLVLGWOAPPGARSMTPCSCGSDLYLVTRHADVIPVRR	1145
DB	61	GSXTLAGPKGIPTQMTYNVDQDLVGWOAPPGARSLTPTCTCGSDLYLVTRHADVIPVRR	120
QY	1146	GDSRGSLLSPRPVSYLKGSGGPLLCPSHGVGVFRAAVCTRGVAKAVDFIPVESMETTM	1205
DB	121	GDSRGSLLSPRPVSYLKGSGGPLLCPSHGAVGIFRAAVCTRGVAKAVDFIPVESMETTM	180
QY	1206	RSPVFTDNSTPPAVPQTFQVAHLHAPTGSGKSTKPAAYAAQGYKVLVLPNSVAATLGF	1265
DB	181	RSPVFTDNSTPPAVPQTFQVAHLHAPTGSGKSTKPAAYAAQGYKVLVLPNSVAATLGF	240
QY	1266	AYMSKAGIDPNRTGVRTITTTGGSITYTYGKFLADGGCGSGAYDIIICDECHSTDSTT	1325
DB	241	AYMSKAGIDPNRTGVRTITTTGAPITYTYGKFLADGGCGSGAYDIIICDECHSTDSTT	300
QY	1326	ILGIGTVLQAEATAGARLVLATATPPGCVTVPHNPTEIEIGLSNGBIPFYGKAIPETAI	1385
DB	301	ILGIGTVLQAEATAGARLVLATATPPGCVTVPHNPTEIEEVALSSTGBIPFYGKAIPETI	360
QY	1386	KGGRHLIFCHSKKFCDELAALKJGLGNAYAYRGLDVSUVPPTGDVVVATDALMTGFT	1445
DB	361	KGGRHLIFCHSKKFCDELAALKJGLGNAYAYRGLDVSUVPPTSGDVIIVATDALMTGFT	420
QY	1446	GDPSVIDCNTVQTVDFSLDPTFTTETTTVPQDAVSRSORRGTRGSGIYRFVTPG	1505
DB	421	GDPSVIDCNTVQTVDFSLDPTFTTETTTVPQDAVSRSORRGTRGSGIYRFVTPG	480
QY	1506	ERPSGMFDSSVLCYCDAGCAWYELTPAETSVRLRAYLNTPLPVCODHLEFWSVFTGL	1565
DB	481	ERPSGMFDSSVLCYCDAGCAWYELTPAETSVRLRAYLNTPLPVCODHLEFWSVFTGL	540
QY	1566	THIDAHFLSQTQKAGDNPYLVAYQATVCARAQAPPSPSDQMWKCLIRKLTPLHGPTLL	1625
DB	541	THIDAHFLSQTQKAGDNPYLVAYQATVCARAQAPPSPSDQMWKCLIRKLTPLHGPTLL	600
QY	1626	YRLGAVQNEVILTHPITKYIIMACMSADLEVTVSTWLVGGVLAALAAAYCLITGSGVIVGR	1685
DB	601	YRLGAVQNEVITTHPITKYIIMACMSADLEVTVSTWLVGGVLAALAAAYCLITGSGVIVGR	660
QY	1686	IILSGKPAVDPREVLYQEFDEMESCASQLPYIEQGMQLAQFKOKAIGLLQATKQAEA	1745
DB	661	IILSGKPAIIPDREVLYREFDEMESCASHLPYIEQGMQLAQFKOKAIGLLQATKQAEA	720
QY	1746	AAPVVEKWRLETFWAKHMNFISGIOYLAGSLTLPGNPAIASLMAFTASITSLPTTON	1805
DB	721	AAPVVEKWRILEAFWAKHMNFISGIOYLAGSLTLPGNPAIASLMAFTASITSLPTTOH	780
QY	1806	TLLFNILGWWAAQLAPPSSAASAFVAGIAGAGVSGIGLGVLDILAGYGAGVAGALVA	1865
DB	781	TLLFNILGWWAAQLAPPSSAASAFVAGIAGAGVSGIGLGVLDILAGYGAGVAGALVA	840
QY	1866	FKVMSGEVPSBDBLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQMNPLIAFASR	1925
DB	841	FKVMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQMNPLIAFASR	900
QY	1926	GNHVSPTHVPESDAAARVTOILLSSITITQLKRLHQWINEDCSTPCSGSLDRDVMWDIC	1985
DB	901	GNHVSPTHVPESDAAARVTOILLSSITITQLKRLHQWINEDCSTPCSGSLDRDVMWDIC	960
QY	1986	TVLTDFTKWLQSKLLPRLPGVPPFLSCORGKGYKGYVRGDMQITTCPCGAQIAGHVXGSMR	2045
DB	961	TVLTDFTKWLQSKLLPRLPGVPPFLSCORGKGYKGYVRGDMQITTCPCGAQIAGHVXGSMR	2100
QY	2046	IVGPTCSNTWGTTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGFHYVTGM	2105
DB	1021	IVGPTCSNTWGTTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGFHYVTGM	2180
QY	2106	TTDNVVKPCQVPAPEFTEVDGVLHRYAPACKPLAREVTFVQGLNOYLVGSQLCPEPE	2165
DB	1081	TTDNVVKPCQVPAPEFTEVDGVLHRYAPACKPLAREVTFVQGLNOYLVGSQLCPEPE	2140
QY	2166	PDVTVLTSMLTDPSSHITAEAKRLARGSPPLASSASQSLAPSLKATCTTHHDSPDAD	2225
DB	1141	PDVAVLTSMLTDPSSHITAEAKRLARGSPPLASSASQSLAPSLKATCTTHHDSPDAD	2200
QY	2226	LIBANLLWRQEMGNITRVESENKVVILDSFEPLHABGDEREISVAAEILRKSRKFPASAL	2285
DB	1201	LIBANLLWRQEMGNITRVESENKVVILDSFEPLHABGDEREISVPAEILRRSRKFPAM	1260
QY	2286	PIWAPDYNPLLESWKPDYVPPVHGCPLPPTKAPPIPPRRKRTVIVTESNVSSALA	2345
DB	1261	PIWAPDYNPLLESWKPDYVPPVHGCPLPPTKAPPIPPRRKRTVIVSESTVSSALA	1320
QY	2346	ELATKTFGSSGSAVDSCTATALPDLASDDGKSDVSESYSSMPPLGEGCDPLSDGSM	2405
DB	1321	ELATKTFGSSGSAVDSCTATAPDQPSDDGKSDVSESYSSMPPLGEGCDPLSDGSM	1380
QY	2406	STVSEASEDVCSSMSYTTWGALITPCAABESKLPINPLNSLLRHHNNVYATTSRSAS	2465
DB	1381	STVSEASEDVCSSMSYTTWGALITPCAABETKLPINALNSLLRHHNNVYATTSRSAS	1440
QY	2466	LRQKKVTFDLRLQVLDHYRDVLKEMKAKASTVAKALLSIEACKLTPPHSAKSKFGYAK	2525
DB	1441	LRQKKVTFDLRLQVLDHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFGYAK	1500
QY	2526	DVRNLSRAVNHIRSVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLIVFPDLG	2585
DB	1501	DVRNLSKAVNHIRSVWKDLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560
QY	2586	VRUCEKVALYDVVSTLPOAVNGSSYGFQYSPQORVFLVNTWKSCKPMGFSYDTRCFDS	2645
DB	1561	VRUCEKVALYDVVSTLPOAVNGSSYGFQYSPQORVFLVNAWKAKCKPMGFAYDTRCFDS	1620
QY	2646	TVTESDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLNSKGNGCYRRCRASGVLT	2705
DB	1621	TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLNSKGNGCYRRCRASGVLT	1680
QY	2706	TSCGNTLTCLYUKATAACRAAKLQDCTMLVNGDDLVLWICESAGTQEDAAALRAFTAMTRY	2765
DB	1681	TSCGNTLTCLYUKAAAAACRAAKLQDCTMLVNGDDLVLWICESAGTQEDDEASLRAFTAMTRY	1740
QY	2766	SAPPDPPQPEYDILELITSCSSNVSVAHDSAGKRVYVLTTRDPTTPLASAAWETAHTPTIN	2825
DB	1741	SAPGDPKPEYDILELITSCSSNVSVAHDSAGKRVYVLTTRDPTTPLASAAWETAHTPTIN	1800
QY	2826	SWLGNIIWYAPTLWARMILMTHFFSILLAQLEKALDCQIYGACYSIEPDLPIIIRL	2885
DB	1801	SWLGNIIWYAPTLWARMILMTHFFSILLAQLEKALDCQIYGACYSIEPDLPIIIRL	1860
QY	2886	HGLSAFTLHVSPEINRVASCLKLGVPPLRTWHRARSVRALKLSGGGAATCGRYLF	2945
DB	1861	HGLSAFTLHVSPEINRVASCLKLGVPPLRTWHRARSVRALKLSGGGAATCGRYLF	1920

[illegible]

```

Qy 713 AIKWEYILLFLLLADARVCACLMMWLLIAQAEAALENVLVNLNAAASVAGHGILSFLVFF 772
Db 541 AIKWEYVVLFLLLADARVCSCLLMMWLLISQAEAALENVLNLAASLAGTHGLVSFLVFF 600

Qy 773 CAANYIKRIAPGAAAYAFYGVWPLLLLILLALPPRAYA 809
Db 601 CFAWYLGKRWPGAVYTFYGMWPLLLLILLALPPRAYA 637

RESULT 4
US-11-117-667-2
; Sequence 2, Application US/11117667
; Publication No. US20050260567A1
; GENERAL INFORMATION:
; APPLICANT: Roche Palo Alto LLC
; APPLICANT: Acrill, Andrew
; APPLICANT: Orle, Karina
; APPLICANT: Paterson, Morris
; TITLE OF INVENTION: NS5A Nucleotide Sequence Variation As A Marker For Interferon
; TITLE OF INVENTION: Response
; FILE REFERENCE: R0216B-REG
; CURRENT APPLICATION NUMBER: US/11/117,667
; CURRENT FILING DATE: 2005-04-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 446
; TYPE: PRt
; ORGANISM: Hepatitis C virus
US-11-117-667-2

```

RESULT 5
US-10-714-781A-67
; Sequence 67, Application US/10714781A
; Publication No. US20050255127A1

GENERAL INFORMATION:

APPLICANT: LOOMORE, SHEENA MAY
 TITLE OF INVENTION: RECOMBINANT VACCINE AGAINST WEST NILE VIRUS
 FILE REFERENCE: 574313-3161.5
 CURRENT APPLICATION NUMBER: US/10/714,781A
 CURRENT FILING DATE: 2003-11-17
 PRIOR APPLICATION NUMBER: 10/679,520
 PRIOR FILING DATE: 2003-10-06
 PRIOR APPLICATION NUMBER: 10/374,953
 PRIOR FILING DATE: 2003-02-26
 PRIOR APPLICATION NUMBER: 10/116,298
 PRIOR FILING DATE: 2002-04-04
 PRIOR APPLICATION NUMBER: 60/281,923
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 10/676,502
 PRIOR FILING DATE: 2003-09-30
 PRIOR APPLICATION NUMBER: PCT/FR02/01200
 PRIOR FILING DATE: 2002-04-05
 NUMBER OF SEQ ID NOS: 85
 SOFTWARE: PatentIn Ver. 3.2
 SEQ ID NO 67
 LENGTH: 3433
 TYPE: PRN
 ORGANISM: West Nile virus
 US-10-714-781A-67

Query Match 1.4%; Score 218; DB 6; Length 3433;

Best Local Similarity 19.3%; Pred. No. 5.2e+04;
 Matches 342; Conservative 219; Mismatches 647; Indels 566; Gaps 84;

515 VGGTDRSGVPTVS-----WGENETDVMLANNT-----PPQGNWFGCTWMNS-----TGFT 561
 974 IIGTAVKNLAHSDLSYIESR-----LNDTWKLERAVILGEVKSCTWPETHLWGGIL 1028
 562 K-----TCGGP-----PCNIGGVGNRTLICP-----TDCFRKHP 591
 1029 ESDLIPVTLAGPRSNRRRPGYKTONQGWDEGRVIDFDYCGTIVTILSESCHRG 1087
 592 ATYTKCGSGPWLTPRCLVDYPYRLMHPCTLNFSIFKRMVYGVGVEHRLNAACNW-----T 647
 1088 ATRTTESGKLITDWC-----RSCIL-----PPLRYQTDGSC-WYGWEI 1126
 648 RGER-----CNLEDRDRSELPSLLSTTWQILPCAFT---TLPALSTGLIH 691
 1127 RPQRHDEKTLVQSQVNAYNADMIDPQLGLLVFLQATVLEVRKRWAKISMPAILIALV 1186
 692 L-----HONIVVOYLYGVGSAP-----VSPAI-----KW-- 716
 1187 LVFGGITVTDV--LRYVILVGAFAESNSGGDVVHLALMATFKIQPVFMVASFLKARWTN 1244
 717 -EYILL-----FLLLA--DARVCACLMMLLIQAEEALENLVLNAAASVAGAHGILSFL 769
 1245 QENTILLMAAVFFOMAYHDAQO-ILLW-----EIPDVLSLAVAWM--ILRAI 1289
 770 VFCAAWYIKGLA-----PGAAYAFYGVWPLLLILLALPPRAYALDRE---MAASCGGAVL 823
 1290 TTTTTSNVVVPLALLTGLCLNLDVYRILLMVIG----SLIREKRSAAKKKGASL 1345
 824 VGLVFLTSLPYKVFELRLIWMLOYFITRAEAMQVWVPPNLNVRGRDAIILLTCVHPE 883
 1346 LCLALASTGLFNPMLA-----AGLIACDP--NRKRGWPATEVMTA----- 1384
 884 LIFDITKLLAILGLBMLVQAGITRVPYFVRAQGLIRACMLVRKVAGGHYVQ-----WVFM 939
 1385 -----VGLMFAIVGLABLDIDSMAPMTI--AGLMFAAFVISGKSTDMWERTADISWE 1437
 940 KLGALTGT-----YVYNHLTPLEDWAHAGRLAVAVEPVV-FSAMETKVI 984
 1438 SDAEITGSSERVDRLDDGNFQLMN--DPGAPWKIWMRLMVCLAISAYTMAILPSVVG 1495
 985 TWGADTAACGDIILGLPVSARRG-----KEIFLGPADSLGCGWRLLAPITAYSQOT 1036

1496 FW-----ITLQYTKRGVLWDTPSPKEYKKG--DTTTG-----VYRIMT 1532
 1037 RGVLCIITSLTGRDNQVEGEVQVSTATQFLATCINGVCTVYHAGSKTLACPKGP 1096
 1533 RGLLG-----SYQAGAGVMVEGVFHTLWHHTTKGAALMSGEGR 1569
 1097 ITQMYTNVDLDLVGWOAPFGASMTPCSCGSDLYLVTRHADVIP--VRRRGDSRGLL 1153
 1570 LDPYMGSKVEDRLCYGGPKLQHKWN---GQDEVQWIV-----VEPGKNVKNVOTKGVFK 1622
 1154 SPR-----PVSYLKSGSGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVESMETMRS 1207
 1623 TFEIGEIGAVTLDPPTGTSPIVDKNGDVI GLYGVINGVIMPGSYSAIVQGERMDRPIA 1682
 1208 PVFTDNSTPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVNLPSVAATLGFAY 1267
 1683 -----GFEPEMLRKQITVLDLHPGAGKTRRLPQIIEKAINRRLRTAVLAPTRVVAE 1736
 1268 MSKA-HGIDPNIRTVRTITGSSITYSYTYGKFLADGGCSGA-----DYIII 1314
 1737 MAEALRGLPIRYOT-----SAVPREHNGNEIVDMVCHATLTHRLMSPHRPNVNLV 1788
 1315 CDECHSTDSITLIGTVDLQAEATAGARLVLATATPPGSVTVPHPNIEIGLSNNGEIP 1374
 1789 MDEAFTDPASIAARGYSTKVB-LGEAAAI FMTAIPPGT-SDFFP-----ESNSPIS 1839
 1375 FYCKAIPIEAKGGRHLI-----FCHSKKKCDELAALKTGLGLNAVAYYRGLDVSV 1425
 1840 DLQTEIPDRRAWNGYEWIETGKTWVFPVSVKMGNEIALCLQORAGKVKVQLNKRKSYETE 1899
 1426 IPPIGD---VVVVATDALMTGTFDSDVIDCNTVTQTVDFSLDPTFTI-ETTTVPQDA 1481
 1900 YPKCKNDMDPFTVITDISEMGANFKASRVDSRKSVKPTIITEGEGRVILGEPSAV--TA 1957
 1482 VRSQRGRGTRGRSGIYRFTVEGERSGMPDSVLCY-----DAGCAWELTPAE 1534
 1958 ASAAQRRGRIGR-----NPSQVGD-----YCYGGHTNEDDSNFAHW----- 1994
 1535 TSVRLRA-YLNTPLPVCODHLEFWESVFTGLTHIDA-----HFLSQTQKQAGDNF 1583
 1995 TEARIMLDNINMGGLIAQFYQPEREKYV-----MDGEYRLRGEERKNFLELLRTA--DL 2048
 1584 PVLAVAYQATCARAQAPPPSWDQWKKLIRLKPETHGPTPLLVRLGAV--QNEVILTHP 1640
 2049 PVLAVAYV-----AAAGSVHDDRW-----CFDGP-----RTNILEDNNEV---EV 2087
 1641 ITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSSVIVGRIILSGKPAVVPREV 1700
 2088 ITKL-----GERKILRPMI-----DARVYSDHQA 2112
 1701 LYQFDEMEBECASQLPVIEQGMQLAEOPKQKALGLLOTATYKQAEAAAPVVESKWRALETF 1760
 2113 LKAFKDFASGKRSQIGLIEVLGKMPHEFMGKT-----WEALDTM 2151
 1761 W-----AKHWNFISGQYLAGLSTLPGNPAIASIAMAFTASITSPITQNTLLENILGCV 1816
 2152 YVWATAEK-----CGRAHRMALEELPDALQTIALLISVMT----- 2188
 1817 AAQLAPPSAASAFVAGIAGAAVGSIGLKGVLVDILAGYAGVAGALVAFKVMGSEVPST 1876
 2189 -----MGVFFLLMORKGIGKI-----GLGGAVLGVATFFCWMMA-EVPGT 2226
 1877 EDLVNLLPAIL-----SPGALVVGWCAAILRRHVRHVGPOGEGAVQWNRILIA 1921
 2227 KIAGMLLSLLMIVLPIPEPEKORSQTDNQLAVFLICVMTLVSAVAANE--MGWLDKTKS 2284
 1922 FAS-----RGNHVSPTHYVPESDAAARVQTILSSLTITQLLKRL--HOWINEDCS- 1969
 2285 DISSLFGQRIEVENKFNFSMGEFLDLRPAATANSVAVTTAVLTPL/KHLITSDYINTSLTS 2344
 1970 -----TPCSSLWLDVWDWICTVLTDFKTLWQSKLLPRLPGVPFLSCQGV----- 2015
 2345 INVQASALFTLARGFPFVDV--GVSALLAAGCGQVTLTVTVAATLLFCHYAYWPGW 2402

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 6
 US-10-821-234-1055
 ; Sequence 1055, Application US/10821234
 ; Publication No. US20050255114A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Labat, Ivan
 ; APPLICANT: Stache-Crain, Birgit
 ; APPLICANT: Andarmani, Susan
 ; APPLICANT: Tang, Y. Tom
 ; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
 ; FILE REFERENCE: 821A
 ; CURRENT APPLICATION NUMBER: US/10/821,234
 ; CURRENT FILING DATE: 2004-04-07
 ; PRIOR APPLICATION NUMBER: US 60/462,047
 ; PRIOR FILING DATE: 2003-04-07
 ; NUMBER OF SEQ ID NOS: 1704
 ; SOFTWARE: pc_seq_genes Version 1.0
 ; SEQ ID NO 1055
 ; LENGTH: 1043
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-821-234-1055

Query Match 1.1%; Score 171; DB 6; Length 1043;
 Best Local Similarity 23.3%; Pred. No. 0.0002;
 Matches 96; Conservative 53; Mismatches 141; Indels 122; Gaps 19;

Qy 1199 ESMETTWES-VFTDNTSTPPAVPQTFQVHLHLAHTGSGKSTKVP-----AAYAAQGVKVL 1252
 Db 392 ESIQAVRKSLPVFPFRELLAAIANHQVLIIEGTGSGKTTQIQYLFEEGYTNKGMKIA 451

Qy 1253 VLNPSSVAATLFGAYMSKAHG-----IDPNIR-----TGVRTITGGGSIYTVG-----K 1298
 Db 452 CTQPRVAVMSVAARVAREMGVKNLNEGVYSIRPEDCTSERIV-----LRVMTDGMLLRE 506

Qy 1299 FLADGCGGAYDIIICDECHSTDSITLIGTGLDQAEATAGARLVVLATATP----- 1351
 Db 507 FLSEPDLA--SYSVVMVDEAHERLTHDILFGLIKDVARPRPELKVIVASATMTARFST 564

Qy 1352 -----PGSVTVPHNTEELGSLNNGEIP---FYGKAIPAEATKG----- 1387
 Db 565 FDDAPVPRIFGR-----RFPVDIFYTKAPEADYLEACVSVLQIHVT 607

Qy 1388 ---GRHLIFCHSKK---CDELAAKLTGLG---LNAVAYYRGLDVSV-----IPP 1428
 Db 608 QPQGDILVLTGQEEIEAECMLQDRCRRLGSKIRELLVLVIYANLPDMDQARIFQPTTP 667

Qy 1429 IGDVVVVATDALMTGFTGD-FDSVIDNCNTCTVTQTVDFSLDPTFTIETTIV-PQDAVRSRQ 1486
 Db 668 GARKVVVATNIAETSLTIEGIYVLDPCFKQK---SYNPRTGWESLITVTPCASKANQ 723

Qy 1487 RRGRTGRGRGIYFVTPGERPSGMFSDSSVLCECYDAGCAW-YELTFAETSV 1537
 Db 724 RAGRAGRVAAG-----KCFRLYTAWAYQHELEETTIV 754

RESULT 7
 US-10-949-720-389
 ; Sequence 389, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 8
 US-10-949-720-419
 ; Sequence 419, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 9
 US-10-949-720-432
 ; Sequence 432, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 10
 US-10-949-720-454
 ; Sequence 454, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 11
 US-10-949-720-476
 ; Sequence 476, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 12
 US-10-949-720-498
 ; Sequence 498, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 13
 US-10-949-720-520
 ; Sequence 520, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 14
 US-10-949-720-542
 ; Sequence 542, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 15
 US-10-949-720-564
 ; Sequence 564, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 16
 US-10-949-720-586
 ; Sequence 586, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 17
 US-10-949-720-608
 ; Sequence 608, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 18
 US-10-949-720-630
 ; Sequence 630, Application US/10949720
 ; Publication No. US20050249736A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Krasnoperov, Valery
 ; APPLICANT: Zozulya, Sergey
 ; APPLICANT: Kertes, Nathalie
 ; APPLICANT: Reddy, Ramachandra
 ; APPLICANT: Gill, Parkash
 ; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

Qy 2016 -----KGWRG-----DGIMQTTCP 2030
 Db 2403 QAEAMSAQRRTAAAGIMGNVAVDGVATDVP 2433

RESULT 19
 US-10-949-720-652
 ; Sequence 652, Application US/10949720
 ; Publication No.

```
FILE REFERENCE: VASG-P02-002
CURRENT APPLICATION NUMBER: US/10/949,720
CURRENT FILING DATE: 2004-09-23
PRIOR APPLICATION NUMBER: US 60/454,432
PRIOR FILING DATE: 2003-03-12
PRIOR APPLICATION NUMBER: US 60/454,300
PRIOR FILING DATE: 2003-03-12
PRIOR APPLICATION NUMBER: US 10/800,350
PRIOR FILING DATE: 2004-03-12
NUMBER OF SEQ ID NOS: 425
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 419
LENGTH: 401
TYPE: PRT
ORGANISM: Homo sapiens
US-10-949-720-419

Query Match 0.8%; Score 131; DB 6; Length 401;
Best Local Similarity 22.5%; Pred. No. 0.034;
Matches 85; Conservative 39; Mismatches 125; Indels 128; Gaps 19;

QY 1940 AARVQTILSSLTITOL--LKLHWINEDCSTPCSGSMLRDVMDICTVLDTFKWLQ 1996
Db 13 AALAEETLLNTKUEQTQTLVNLTRPPETVPRELVPVAVAGSCVDA----- 56

QY 1997 SKLLPRLPGVPFLSCORGYKGVWRGDGIMQTTCPGAGIA-GHVKNGSMRIVGPRTCNT 2055
Db 57 --VPAPGSPSLYCRE--DGWAEQPVTCSCAPGFEAAEGNTK-----CRAC 100

QY 2056 WHGTF-PINAYTTGCTPSPAPNVSRALMRVAEEYVEVTRVGDPHVVTGTTDVKPC 2114
Db 101 AQGTGFKPLSG--EGSCQPCPANSNHTIGSAVCO-----CRVGYFR----ARTDPRGAPC 149

QY 2115 QVP--AP-EFTEVDGVRLH-RYAPACKPLLRDVTQVGLNQYLV-GSQLPC-----E 2163
Db 150 TTPSAPRSVSRVRLNGSLHLEWAPLESGGREDLTVALRCRCRPGSCAPCGGDLTFD 209

QY 2164 P-----EPDVTV--LTSMLTDPSHITAETAKRELARGSPPLSSASSASQLSAPSLKATC 2215
Db 210 PGPRDLVEPVVVVGLRPDPFTYTFVTALNGVSSLATGPV----- 250

QY 2216 TTHDSDPADLIEANLLWROEMGNITRVESENKVILDSPEPLHAEGDEREISVAEIL 2275
Db 251 -----FEPVNVITD-REVPPAVSDI 269

QY 2276 RSKRFPSPALPI-WARP 2291
Db 270 RVTRSSPSLSLAWAVP 286

RESULT 9
US-11-108-172-1116
; Sequence 1116, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon B.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary A. W.
; APPLICANT: Vedwick Thomas S.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.471C15
CURRENT APPLICATION NUMBER: US/11/108,172
CURRENT FILING DATE: 2005-04-15
PRIOR APPLICATION NUMBER: US 10/025,380
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 09/922,217
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 09/833,263
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: US 09/649,811
PRIOR FILING DATE: 2000-08-28
PRIOR APPLICATION NUMBER: US 09/609,448
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/575,251
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/519,444
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: US 09/504,629
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: US 09/480,321
PRIOR FILING DATE: 2000-01-10
PRIOR APPLICATION NUMBER: US 09/476,296
PRIOR FILING DATE: 1999-12-30
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1130
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1116
LENGTH: 5405
TYPE: PRT
ORGANISM: Homo sapiens
US-11-108-172-1116

Query Match 0.8%; Score 130; DB 7; Length 5405;
Best Local Similarity 17.8%; Pred. No. 2.6;
Matches 628; Conservative 322; Mismatches 1160; Indels 1426; Gaps 181;

QY 104 RPSGPTDPPRRSRLKVIDTLTCG-----FADLMGYIPLVGAPLGA 147
Db 300 RPSM-----PLYLSANVG--IQVLLFGTGAIRNEVTPYLVLPDVAAYCPA----- 345

QY 148 ARALAHGVRVLEDGVNATGNLPGCSFISILLA---LLSCITI-----PAS 190
Db 346 -----YVKSVPCEGVALVVAQTKAISGLTIDGHAVGAKLTWEAVPGS 389

QY 191 AYEVRNY-----SGIYHVTDNCSN-----SSIVYEAADVIMHT-----PGC----- 226
Db 390 EFSYAEVELGTADMIHTAEATNLGLTLFGLAKAIGYATAADCGRTVLSPVPESCGCMQC 449

QY 227 ---VPCVQEGNSRCWALTFTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYDGLCGS 283
Db 450 AAGQRCQVVGKAGCAESTAVCRAQGDPHYTTTDFGRRYDM-GT---CSYTMV-ELCSE 504

QY 284 IFLYSQLFTES-----PRRHETVQDCNCSIYPGHVSGHRMADMMNWSPTLWV 334
Db 505 ---DDTLPAFSVEAKNHRGSRVSYVGLTVTRAYSHSVLTRGEVGVFVVDNQSRPLV 561

QY 335 S---QLLRI-----PQAVVDMVAGAHWGLAGLAYYSVMGNWAKVLI VALLFAGVDGETHT 387
Db 562 SLSEGRLEVYQSGPRAVVELVFG-----LVVYTDWDCQLALSPLARFQDQVCGGL 610

QY 388 TGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSWHINRTALNCNDSLQGTGTFALPYAKHF 447
Db 611 CGNYNGDPADDFLTDPGALAPDAVEFAS---SWKLDGDGYLDCED----- 651

QY 448 NSSCCPERMASCRIIDWFQAQGWGITTYTKPNSSDORPCYWHY-APRPGGVVPASQVCGP- 505
Db 652 ---GCNNCPACTP-----GQAQHYEGDRLCGML--TKLDGPF 684

QY 506 VYC---FTPSP-----VVVGTDRS---GVPTYSGENETDMLNLTNTPPQGNWF 550
Db 685 AVCHDTLDRPFEQCQVYDLGVGGERLSLRCGLSAYACALGLGIS-VGDWRSRSPANCPL 743

QY 551 GC-----TWMNSTGTCTCGPPCNCITGGVGNRTLIC-PTDCFRKHPEATYT 595
```


QY 2414 EDVVCMSYTTWTGALITP-----CAAERSKLP1-----NPLNSLLR----- 2451
Db 2368 RECVCDAAGVLSGDTCPVVGCGCLHDDRYYPPLGQTFYPPGCGDCLCRREGVSCBP 2427
QY 2452 -----HHN-----MYATTSSASLRKKVTFDLQVLVDHHRDLVLMKEMAKAS 2495
Db 2428 SSCGPETHCRPSGSGSLGCVAVGTTTQASG-----DPHY-----T 2462
QY 2496 TVKAKLLSTEEACKLTPHSAKSK-----FYGAKDV-RNLSSRAVNH1-- 2538
Db 2463 TPDGRFPDMGTCTVYVLAQTCGRPLGLHFAVLQENVANGNRVSVTRVITVQVANFTLR 2522
QY 2539 --RSVNEDLLE--DTPTIDTTIMAKSEVCPQEPKGRKPARLIIVPDLGVRVCEKMAL 2594
Db 2523 LEQRMKVTVNGVDMKLPV-----VLANGQI-----RASQHSQDWIETDFGLRVAYDLVY 2573
QY 2595 YDVVSTLP-----QAVGSSYGFQYSPKQ-----RVEFLVNTWKS----- 2629
Db 2574 Y-VRVTVPGNYIQLMCGLCGNYNGDPKDDFKPQNGSQAGNANBFGNSWEEVVPDSCPLPP 2632
QY 2630 KKCPCMGFSYTRCFDSTVTESDIRVEESIIYQCCDLAPEARQAIRSLTERLYIGGPLTNSK 2689
Db 2633 PTCPPG-----SEGCIPSECEPPELEKKYKEEFCGLLSSPTGPLSSCHKLVDPQGPL----- 2685
QY 2690 GQNGCYRRCRASGVLTTSCGNTLTCLYKATAA----- 2721
Db 2686 -KDCIPDLCLGGNLSILCSN-IHAYVSAQMAAGHVPEWRNETFCPMECPQNSHVELCA 2743
QY 2722 -----CRA--AKLQ-----DCTMLVNGDDLVI-----CESAGT---QEDAAAL 2755
Db 2744 DTCSLGCSALSAPLQCPDCAEGCQCDGFLYNGQACVPIQCCGCHNGAYRPEQTVLI 2803
QY 2756 RATEAMT-----RYSAPPDPPQPEYDLELITSCSSNVSAHADSGRKYVYLTR 2805
Db 2804 DNCRQOCTCHAGKVVVCOEHSCKRPGVCQP----- 2843
QY 2806 DPTTPPLAARAAWETARHTPINSWLGN-----IMYAPTILWARMILMTHFFSILLAQOLEKA 2861
Db 2844 DPHGVTCRQETCKEQ-----GGQGVCLPNYEATCMLMGDPHYHSGFD-----GRK 2889
QY 2862 LDCQIYGAC-YSIEPLDLPOIIEHLGLSAPTLHSYSGEINRVASCLRLKGVPPPLRT-- 2918
Db 2890 FDFQ--GTCNVLATTCGPGV--STQGLTPFTVTTKNQNRGNPAVSVYRVVTVAAALGTNI 2945
QY 2919 -----WRHARSVRAKL-----LSQGGRAATC-----GRYLFNWAIVTKLK 2954
Db 2946 SIHKDEIGKVRVNGVLTALPVSADGRISVAQGAASKALLVADFLQVSYDNMNRVDVTL- 3004
QY 2955 LTPIPAASQLDLSGMFVAG-----YSGGDIYHSL-----SRARPRWFPLC 2994
Db 3005 ----PSSYHGAVCG--LCGNMDRPNNDQVFPNGTLAPSIPIWGGSWRAPGWDPLC 3054

RESULT 10

US-10-821-234-1120
; Sequence 1120, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt seq_genes Version 1.0
; SEQ ID NO 1120
; LENGTH: 4384

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1120

Query Match 0.8%; Score 129; DB 6; Length 4384;

Best Local Similarity 20.6%; Pred. No. 2.2;

Matches 178; Conservative 93; Mismatches 305; Indels 288; Gaps 43;

QY 2049 PRTCS-NWHGTFPINAYTTGPCT-PSPA-----PNYS--RALMRVAEEY 2090
Db 1496 PTTSYKPFSTPRYQSWTTAPITVGPAGKSTLSSSSSNTSPASPLKSIWSVSTPSP 1555
QY 2091 VEVTRVGDPHYVTGMTDNNVKPCQVPAP-----EFTVEVDGVRLLHR 2132
Db 1556 IKST-----LCASTTSSVKSISDVASPIRSLRTMSSPIKTVVQSYPNIQVSSGTLAR 1608
QY 2133 YAPA-----CKPL--LRBDVTFOVLNQYLVGSQL-----PCPEPDVTVLTSM LTD 2177
Db 1609 -APAVTEATPLKGLASNSTFSSRTSPVTTAGSLERSSITMTTPASPKNINMYSSSLPF 1667
QY 2178 PSHITAETAKRLARGSPFSLASSASQLSAPSLKATCTTHHDSPPADLLEANLLWRQEM 2237
Db 1668 KSIIT-----SAAPLISSP-LKSVSPVKSIRDV--ISSAKITMASSL 1707
QY 2238 GGNITRVESENKVILD-SPEPLHAGDEREISVAAEILKRKRKFPFSALEPIWARPDYNPP 2296
Db 1708 SSPVKQMPGHAELVNGSISPL-----KYASS----- 1735
QY 2297 LLESWKDPDYVPVHVHGCPLPPTKAPP1PPRRKRRTVVLTVESNVSSALAEATKTFGSG 2356
Db 1736 -----STLINGC-----KATATLQE-KISS-----ATNSVSVV 1763
QY 2357 SSAVDS-----GTATALPDLASDDGDKGSDVESYSSMPPLEGEPDGLDGGSWSTVSEE 2411
Db 1764 SAATDTVEKVFSTTTAMP-----FSPLSYVSAAPSAFO-----SLRTSPAS 1805
QY 2412 ASEDVCCSSYTW--TGALIT-PCAAEESKLPIN-----PLNSLLRHHNMVYATTSR 2462
Db 1806 ALYTSLGSSISATSSVTSSIIIVPVYVVNVLPPEPALKKLPDSNF-----TKS 1855
QY 2463 SASLRQKKVTFDLQVLDDHYRDLVKEMKAKASTVAKALLSIEBACKLTPPHSAKSKFGY 2522
Db 1856 AALLSPIKLTTHETHPHF-----SRTSPVKSSLPLAPSAKLSTPSSLSS----- 1904
QY 2523 GAKDVRNLSRAVNHIRSVWEDLLETPTIDTTIMAKSEVFCVQPE--KGGRKPARLIV 2580
Db 1905 -----SQEILKDVAEKMDLNR--MTAILQTDVPEEKPF--QPELPKEGR-----I 1946
QY 2581 FPDGLVRVCEKMA--LYDVVSTLPOAVMGSSYGFQYSPKQRFVFLVNTWKSCKCPMGFSY 2638
Db 1947 DDEEPFKIVEKVBEDLVKVEILKKQVCVDNKG---SP-----KSPKSDKGHSP 1992
QY 2639 DTRCFDSTVTESDIRVEESIIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGONCGYRRC 2698
Db 1993 E-----DDWIEFSSEIREARQQA-----AASQSPSLPFRVQVKAKAASEK-----DYNLT 2038
QY 2699 RASGLVTTSCGNTLTCLYKATAACRAAKLQDCTMLVNGDDLVIWICESAGTQEDAAALRAF 2758
Db 2039 KVIDYLINDIGSSSLTNLK-----YKFEDAKKGGGQKRVLKPAIALQEH----- 2084
QY 2759 TEAMTRYAPPDPPQPEYDLELITSCSSNVSV-----AHDA-----SGKRV 2800
Db 2085 -----KLKMPASMRSTSTSEKELCKMADSFEGDTTILESPDDFSQHDQKSP LSDSG--- 2136
QY 2801 YLTRDPTTPPLARAWEETARHTPI 2824
Db 2137 -FETRSEKTPSAPOSAAETTGKPL 2159

RESULT 11

US-10-949-720-425
; Sequence 425, Application US/10949720
; Publication No. US20050249736A1


```
; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Zozulya, Sergey
; APPLICANT: Kertesz, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING
; FILE REFERENCE: VASG-P02-002
; CURRENT APPLICATION NUMBER: US/10/949,720
; PRIOR FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/454,432
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/454,300
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 10/800,350
; PRIOR FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-949-720-425

Query Match      0.8%; Score 127; DB 6; Length 522;
Best Local Similarity 22.2%; Pred. No. 0.11;
Matches 98; Conservative 44; Mismatches 158; Indels 142; Gaps 22;

QY 1882 LLPAILSPGALVGVVCAAILRRHVGPGEVQVWMN-----RLIAFASRGNHVSPTHYVPE 1937
Db 111 LTPAMWENPYIKVDTVAAEHLTRK-RPGAETCKVNVKTLRLGLPSLKAG-----FYLA 163
QY 1938 SDAARVQTQLSSL-----TITQLKRLHQMINEDCSTPCSGSWLRDWDWICTVLDTDF 1991
Db 164 QDQACMALLSLHLFYKKCAQLTVNLTRPPTVPRELVVPVAGSCVDA----- 212
QY 1992 KTWLQSKLLPRPGVPFLSCQRYGKVGWRGDGMOTTCPCGAQIA-GHYKNGSMRIVGPR 2050
Db 213 -----VPAPGPSPLYCRE--DQWAEQPVTCSCAPGFEAAEGNTK----- 252
QY 2051 TCSNTWHGTF-PINAYTTGCTPPSPAPNYSRALMRVAEEYVEVTRVGDHYVTGMTDN 2109
Db 253 -CRACAQGTFFKPLSG--EGSCQPCPANGSHNTIGSAVCQ-----CRVGYFR-----ARTDP 300
QY 2110 VKPCQVQP--AP-EFFTEVDGVRLH-RYAPACKPRLREDVTFQVGLNQYLV-GSOLPC-- 2162
Db 301 RGAPCTTPSPAPRSVVSRLNGSLHLEWSAPLESQGGREDLTVALRCRCRPGSCAPCGG 360
QY 2163 -----EPDVTV--LTSMLTDPHSHTAETAKRRLARGSPPSLASSASQLSAPS 2210
Db 361 DLTFDGPRLDVEPWWVVRGLRPDFTYTFEVTALNGVSSLATGVPV----- 406
QY 2211 LKATCTTHDSDPADLIEANLLWRQEMGNITRVESKNKWLDSFEPLHAEGDEREISV 2270
Db 407 -----PEPVNVTTD-REVPP 420

QY 2271 AAEILKSRKFPSPALPI-WARP 2291
Db 421 AVSDIRVTRSSPSLSLAWAVP 442

RESULT 12
US-10-949-720-424
; Sequence 424, Application US/10949720
; Publication No. US20050249736A1
; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Zozulya, Sergey
; APPLICANT: Kertesz, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Zozulya, Sergey
; APPLICANT: Kertesz, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING

; FILE REFERENCE: VASG-P02-002
; CURRENT APPLICATION NUMBER: US/10/949,720
; PRIOR FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/454,432
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/454,300
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 10/800,350
; PRIOR FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-949-720-424

Query Match      0.8%; Score 127; DB 6; Length 537;
Best Local Similarity 22.2%; Pred. No. 0.11;
Matches 98; Conservative 44; Mismatches 158; Indels 142; Gaps 22;

QY 1882 LLPAILSPGALVGVVCAAILRRHVGPGEVQVWMN-----RLIAFASRGNHVSPTHYVPE 1937
Db 126 LTPAMWENPYIKVDTVAAEHLTRK-RPGAETCKVNVKTLRLGLPSLKAG-----FYLA 178
QY 1938 SDAARVQTQLSSL-----TITQLKRLHQMINEDCSTPCSGSWLRDWDWICTVLDTDF 1991
Db 179 QDQACMALLSLHLFYKKCAQLTVNLTRPPTVPRELVVPVAGSCVDA----- 227
QY 1992 KTWLQSKLLPRPGVPFLSCQRYGKVGWRGDGMOTTCPCGAQIA-GHYKNGSMRIVGPR 2050
Db 228 -----VPAPGPSPLYCRE--DQWAEQPVTCSCAPGFEAAEGNTK----- 267
QY 2051 TCSNTWHGTF-PINAYTTGCTPPSPAPNYSRALMRVAEEYVEVTRVGDHYVTGMTDN 2109
Db 268 -CRACAQGTFFKPLSG--EGSCQPCPANGSHNTIGSAVCQ-----CRVGYFR-----ARTDP 315
QY 2110 VKPCQVQP--AP-EFFTEVDGVRLH-RYAPACKPRLREDVTFQVGLNQYLV-GSOLPC-- 2162
Db 316 RGAPCTTPSPAPRSVVSRLNGSLHLEWSAPLESQGGREDLTVALRCRCRPGSCAPCGG 375
QY 2163 -----EPDVTV--LTSMLTDPHSHTAETAKRRLARGSPPSLASSASQLSAPS 2210
Db 376 DLTFDGPRLDVEPWWVVRGLRPDFTYTFEVTALNGVSSLATGVPV----- 421
QY 2211 LKATCTTHDSDPADLIEANLLWRQEMGNITRVESKNKWLDSFEPLHAEGDEREISV 2270
Db 422 -----PEPVNVTTD-REVPP 435

QY 2271 AAEILKSRKFPSPALPI-WARP 2291
Db 436 AVSDIRVTRSSPSLSLAWAVP 457

RESULT 13
US-10-949-720-387
; Sequence 387, Application US/10949720
; Publication No. US20050249736A1
; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Zozulya, Sergey
; APPLICANT: Kertesz, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING
; FILE REFERENCE: VASG-P02-002
; CURRENT APPLICATION NUMBER: US/10/949,720
; PRIOR FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/454,432
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/454,300
```

; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 10/800,350
; PRIOR FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 387
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Recombinant B4ECv3NT protein
US-10-949-720-387

Query Match 0.8%; Score 127; DB 6; Length 555;
Best Local Similarity 22.2%; Pred. No. 0.11;
Matches 98; Conservative 44; Mismatches 158; Indels 142; Gaps 22;

QY 1882 LPAILSPGALVGVVCAAILRRHVGPGEGAVQWMMN-----RLTAFASRGNHVSPTHVPE 1937
DB 126 LTPAMMENPIKVDTVAAEHLTRK-RPGAETGKVNKTLRLGLPSKAG-----FYLAF 178
QY 1938 SDAARVTOILSSL-----TITQLKRLHOWINEDCSTPCSGSWLRDVMWDMICTVLTFD 1991
DB 179 ODGACMALLSLHLFYKKCAQLTNLTFPPTVPRELVPVAGSCVDA----- 227
QY 1992 KTWLQSKLLPRLPGVFLSCQRYGKVGWRGDGIMQTTCPGQAIA-GHVXNGSMRIVGPR 2050
DB 228 -----VPAPGPSPLCYRE--DGQWAEQPVTVGSCAPGFEAEGNTK----- 267
QY 2051 TCSNTHWGTF-PINAYTTGCTSPAPNYSRALRWVAEEYVEVTRVGDHYVTGTTDN 2109
DB 268 -CRACAQGTGFKPLSG--EGSCQPCPANSHTIGSAVCQ-----CRVGYFR-----ARTDP 315
QY 2110 VKPCQVQV--AP-EFFTEVDGVRHLH-RYAPACKPLLRDVTFOVGLNQYLIV-GSQLPC-- 2162
DB 316 RGAPCTTPSPAPRSVVSRNLGSSHLHLEWSPLESGGREDLTYALRCRCRPGSCAPCGG 375
QY 2163 -----EPDPDVTV--LTSMLTDPHSITAEAKRRLARGSPPLSSASSASQLSAPS 2210
DB 376 DLTFDGPRLDVPFVWVVRGLRDPDFTVTFVTALGVSSLATGPVP----- 421
QY 2211 LKATCTTHDSDPADLIEANLLWRQMGNIITRVESENKVVILDSPEPLHAEGDEREISV 2270
DB 422 -----PEPVNVTTD-REVPP 435
QY 2271 AAEILRKSRKFPSPALPI-WARP 2291
DB 436 AVSDIRVTRSSPSLSLAWAVP 457

RESULT 14
US-10-949-720-386
; Sequence 386, Application US/10949720
; Publication No. US20050249736A1
; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Zozulya, Sergey
; APPLICANT: Kertesz, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING
; FILE OF INVENTION: ANGIOGENESIS AND TUMOR GROWTH
; FILE REFERENCE: VASG-P02-002
; CURRENT APPLICATION NUMBER: US/10/949,720
; PRIOR FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/454,432
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/454,300
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 10/800,350
; PRIOR FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 386
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Recombinant B4ECv3 protein
US-10-949-720-386

Query Match 0.8%; Score 127; DB 6; Length 570;
Best Local Similarity 22.2%; Pred. No. 0.12;
Matches 98; Conservative 44; Mismatches 158; Indels 142; Gaps 22;

QY 1882 LPAILSPGALVGVVCAAILRRHVGPGEGAVQWMMN-----RLTAFASRGNHVSPTHVPE 1937
DB 126 LTPAMMENPIKVDTVAAEHLTRK-RPGAETGKVNKTLRLGLPSKAG-----FYLAF 178
QY 1938 SDAARVTOILSSL-----TITQLKRLHOWINEDCSTPCSGSWLRDVMWDMICTVLTFD 1991
DB 179 ODGACMALLSLHLFYKKCAQLTNLTFPPTVPRELVPVAGSCVDA----- 227
QY 1992 KTWLQSKLLPRLPGVFLSCQRYGKVGWRGDGIMQTTCPGQAIA-GHVXNGSMRIVGPR 2050
DB 228 -----VPAPGPSPLCYRE--DGQWAEQPVTVGSCAPGFEAEGNTK----- 267
QY 2051 TCSNTHWGTF-PINAYTTGCTSPAPNYSRALRWVAEEYVEVTRVGDHYVTGTTDN 2109
DB 268 -CRACAQGTGFKPLSG--EGSCQPCPANSHTIGSAVCQ-----CRVGYFR-----ARTDP 315
QY 2110 VKPCQVQV--AP-EFFTEVDGVRHLH-RYAPACKPLLRDVTFOVGLNQYLIV-GSQLPC-- 2162
DB 316 RGAPCTTPSPAPRSVVSRNLGSSHLHLEWSPLESGGREDLTYALRCRCRPGSCAPCGG 375
QY 2163 -----EPDPDVTV--LTSMLTDPHSITAEAKRRLARGSPPLSSASSASQLSAPS 2210
DB 376 DLTFDGPRLDVPFVWVVRGLRDPDFTVTFVTALGVSSLATGPVP----- 421
QY 2211 LKATCTTHDSDPADLIEANLLWRQMGNIITRVESENKVVILDSPEPLHAEGDEREISV 2270
DB 422 -----PEPVNVTTD-REVPP 435
QY 2271 AAEILRKSRKFPSPALPI-WARP 2291
DB 436 AVSDIRVTRSSPSLSLAWAVP 457

RESULT 15
US-10-949-720-412
; Sequence 412, Application US/10949720
; Publication No. US20050249736A1
; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Zozulya, Sergey
; APPLICANT: Kertesz, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING
; FILE OF INVENTION: ANGIOGENESIS AND TUMOR GROWTH
; FILE REFERENCE: VASG-P02-002
; CURRENT APPLICATION NUMBER: US/10/949,720
; PRIOR FILING DATE: 2004-09-23
; PRIOR APPLICATION NUMBER: US 60/454,432
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/454,300
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 10/800,350
; PRIOR FILING DATE: 2004-03-12
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-949-720-412

```
Query Match      0.8%; Score 127; DB 6; Length 570;
Best Local Similarity 22.2%; Pred No. 0.12;
Matches 98; Conservative 44; Mismatches 158; Indels 142; Gaps 22;

Qy 1882 LLPAILSPGALVGVVCAAILRRHHVGPGEAVQWMN-----RLIAPASRGNHVSPTHYVPE 1937
Db 126 LTPAMMENPYIKVDTVAAEHLTRK-RPGAATGKVNKXTRLRLGPLSKAG-----FYLA 178
Qy 1938 SDAARVTVQILSSL-----TITQLLKRHLHQWINEDCSTPCSGSWLRDVMWICTVLTDF 1991
Db 179 QDQGACMALLSLHLFYKCAQLTVNLTRFETVPRELVPVAVAGSCVDA----- 227
Qy 1992 KTWLQSKLLPRLPGVPFLSLSCORGYKGVWRGDGIMQTTCGQAIA-GHVKNKSMRIVGPR 2050
Db 228 -----VPAPGPSFSLYCRE--DQWAEQPVTCSCAPGEAAEGNTK----- 267
Qy 2051 TCSNTWHGTF-PINAYTTGCTPPSPAPNYSRALMRVAABEVVEVTRVGDHFHYVTGTTDN 2109
Db 268 -CRACAQGTFFKLSG--EGSCQPCPANSHTIGSAVCQ-----CRVGYFR---ARTDP 315
Qy 2110 VKPCQVP--AP-EFFTEVDGVRH-RYAPACKPLLRDVTFOVGLNQYLV-GSOLPC-- 2162
Db 316 RGAPCTTPPSAPRSVVSRLNGSLHLEWSAPLESGRDLTYALRCRCRPGGSCAPCGG 375
Qy 2163 ----EPDVTV--LTSMLTDPSHITAETAKRRLARGSPPSLASSASOLSAPS 2210
Db 376 DLTFDGPRLDVEPWWVVRGLRPDFTYTFEVTALNGVSLATGVP----- 421
Qy 2211 LKATCTTHHSDPADLIEANLLWRQEMGNI TRVESENKVILDSFEPHLAEGDREISV 2270
Db 422 -----FEPVNVTTD-REVPP 435
Qy 2271 AAEILKSRKFPSPALPI-WARP 2291
Db 436 AVSDIRVTRSSPSSLAWAVP 457
```

Search completed: December 3, 2005, 08:49:25
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 04:36:52 ; Search time 1044 Seconds
(without alignments)
16336.881 Million cell updates/sec

Title: US-09-662-454-4

Perfect score: 9595

Sequence: 1 gccagccctctgatggggc.....ggcctctgcagatcatgt 9595

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H/COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PTCUS/COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP/COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE/COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9595	100.0	9595	3	US-09-014-416-4
2	8311.8	86.6	9413	3	US-09-827-688-6
3	8170.4	85.2	9416	2	US-08-324-977-1
4	8170.4	85.2	9416	2	US-08-384-616-1
5	8170.4	85.2	9416	2	US-08-904-686A-1
6	8170.4	85.2	9416	3	US-09-315-850-1
7	8170.4	85.2	9416	3	US-08-823-895A-27
8	8098.2	84.4	9472	3	US-08-150-204E-96
9	7918.6	82.5	11076	3	US-09-539-601-1
10	7901	82.3	11076	3	US-09-539-601-25
11	7899.4	82.3	11076	3	US-09-539-601-19
12	7894.6	82.3	11076	3	US-09-539-601-31
13	7885.8	82.2	9275	3	US-10-259-275-39
14	7809.2	81.4	9030	2	US-08-324-977-13
15	7809.2	81.4	9030	2	US-08-384-616-13
16	7809.2	81.4	9030	3	US-08-904-686A-13
17	7809.2	81.4	9030	3	US-09-315-850-13
18	7201.4	75.1	9599	3	US-09-014-416-6
19	6791.2	70.8	7917	2	US-08-324-977-31
20	6791.2	70.8	7917	2	US-08-384-616-31
21	6791.2	70.8	7917	2	US-08-904-686A-31
22	6791.2	70.8	7917	3	US-09-315-850-31
23	6754.8	70.4	7863	2	US-08-324-977-35
24	6754.8	70.4	7863	2	US-08-384-616-35

ALIGNMENTS

RESULT 1

US-09-014-416-4

; Sequence 4, Application US/09014416

; Patent No. 6153421

; GENERAL INFORMATION:

; APPLICANT: Yanagi, Masayuki

; APPLICANT: Bukh, Jens

; APPLICANT: Emerson, Susanne U.

; APPLICANT: Purcell, Robert H.

; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND

; FILE REFERENCE: 20264276

; CURRENT APPLICATION NUMBER: US/09/014,416

; EARLIER FILING DATE: 1998-01-27

; EARLIER APPLICATION NUMBER: US 60/053,062

; EARLIER FILING DATE: 1997-07-18

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 4

; LENGTH: 9595

; TYPE: DNA

; ORGANISM: Hepatitis C virus

US-09-014-416-4

Query Match	100.0%;	Score 9595;	DB 3;	Length 9595;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 9595;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GCCAGCCCTCTGATGGGGGCGACACTCCACATGAATCACTCCCTGTGAGGAACACTG	60	
Db	1	GCCAGCCCTCTGATGGGGGCGACACTCCACATGAATCACTCCCTGTGAGGAACACTG	60	
Qy	61	TCCTCAGCAGAAAGCGCTAGCCATGCGTGTAGTATGATGCTGCGAGCTCCAGGAC	120	
Db	61	TCCTCAGCAGAAAGCGCTAGCCATGCGTGTAGTATGATGCTGCGAGCTCCAGGAC	120	
Qy	121	CCCCCTCCGGGAGAGCCATAGTGCTGGGAAACCGGTGAGTACACCGGAAATGGCCAG	180	
Db	121	CCCCCTCCGGGAGAGCCATAGTGCTGGGAAACCGGTGAGTACACCGGAAATGGCCAG	180	
Qy	181	GACGACCGGGTCTTTCTTTTGGATCAACCCGCTCAATGCTCGAGATTTGGCGTGCCTCC	240	
Db	181	GACGACCGGGTCTTTCTTTTGGATCAACCCGCTCAATGCTCGAGATTTGGCGTGCCTCC	240	
Qy	241	CGGAGACTGCTAGCGGAGTAGTGTGGTCCGAAAGCCCTTGTGTTACTGCTCATAGG	300	
Db	241	CGGAGACTGCTAGCGGAGTAGTGTGGTCCGAAAGCCCTTGTGTTACTGCTCATAGG	300	

QY 4681 ATTTTGACTCAGTGATCGACTGCAATAATGATGTGTCACCCAGACAGTCGACTTCAGCTTGG 4740
DB |||||
QY 4681 ATTTTGACTCAGTGATCGACTGCAATAATGATGTGTCACCCAGACAGTCGACTTCAGCTTGG 4740
DB |||||
QY 4741 ATCCACCTTACCAATTGAGACGACGACGTCGCCCAAGACGCGGTGTCGGCTCGCAAC 4800
DB |||||
QY 4741 ATCCACCTTACCAATTGAGACGACGACGTCGCCCAAGACGCGGTGTCGGCTCGCAAC 4800
DB |||||
QY 4801 GCGAGGTAGAACTGCGCAGGGTAGAGTGGATCTACAGTTTGTGACTCCAGAGAAC 4860
DB |||||
QY 4801 GCGAGGTAGAACTGCGCAGGGTAGAGTGGATCTACAGTTTGTGACTCCAGAGAAC 4860
DB |||||
QY 4861 GGCCTCGGGCATGTTCCGATCTTCGGTCTCTGTGTAGTGTATGACGCGGCTGTGTT 4920
DB |||||
QY 4861 GGCCTCGGGCATGTTCCGATCTTCGGTCTCTGTGTAGTGTATGACGCGGCTGTGTT 4920
DB |||||
QY 4921 GGTATGAGTCA CGCCCGCTGAGACCTCGGTTAGGTTGCGGCTTACCTAAATACACAG 4980
DB |||||
QY 4921 GGTATGAGTCA CGCCCGCTGAGACCTCGGTTAGGTTGCGGCTTACCTAAATACACAG 4980
DB |||||
QY 4981 GGTGCGCGTCTGCAGAGCACTCTGGAGTTCTGGAGAGGTCTTACAGGCTCACCC 5040
DB |||||
QY 4981 GGTGCGCGTCTGCAGAGCACTCTGGAGTTCTGGAGAGGTCTTACAGGCTCACCC 5040
DB |||||
QY 5041 ACATAGATGCCACATTCCTGTCCAGACTAAACAGCAGGAGACAACTTTCCTTACCTGG 5100
DB |||||
QY 5041 ACATAGATGCCACATTCCTGTCCAGACTAAACAGCAGGAGACAACTTTCCTTACCTGG 5100
DB |||||
QY 5101 TGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCATCGTGGGACCAAA 5160
DB |||||
QY 5101 TGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCATCGTGGGACCAAA 5160
DB |||||
QY 5161 TGTGGAAGTGTCTCATACGCTGAAACCTTACACTGCA CGGGCCAAACACCCCTGCTGTATA 5220
DB |||||
QY 5161 TGTGGAAGTGTCTCATACGCTGAAACCTTACACTGCA CGGGCCAAACACCCCTGCTGTATA 5220
DB |||||
QY 5221 GGTAGGAGCCGTCGCAAAATGAGTTCATCTCACACACCCCACTAAATACATCATGG 5280
DB |||||
QY 5221 GGTAGGAGCCGTCGCAAAATGAGTTCATCTCACACACCCCACTAAATACATCATGG 5280
DB |||||
QY 5281 CATGATGTGCGGTGACCTGAGGTGCTCACTAGCACCTGGTGTGTTAGCGGAGTCC 5340
DB |||||
QY 5281 CATGATGTGCGGTGACCTGAGGTGCTCACTAGCACCTGGTGTGTTAGCGGAGTCC 5340
DB |||||
QY 5341 TTGCACTTTGGCGGATCTGCTGACGACAGGAGGTGTGTTGAGGAGATCA 5400
DB |||||
QY 5341 TTGCACTTTGGCGGATCTGCTGACGACAGGAGGTGTGTTGAGGAGATCA 5400
DB |||||
QY 5401 TCTTGTCCGGGAAGCCAGCTGCTGTCGACAGGAGTCTCTACAGGAGTTCGATG 5460
DB |||||
QY 5401 TCTTGTCCGGGAAGCCAGCTGCTGTCGACAGGAGTCTCTACAGGAGTTCGATG 5460
DB |||||
QY 5461 AGATGGAAGAGTGTGCTTACAACTTTCCTTACATCGAGCAGGAAATGCACTGCGGAGC 5520
DB |||||
QY 5461 AGATGGAAGAGTGTGCTTACAACTTTCCTTACATCGAGCAGGAAATGCACTGCGGAGC 5520
DB |||||
QY 5521 AATTCAAGCAAAAGGCGCTCGGGTTGTGCAAAAGGCGCAACGAGGAGGTGCTG 5580
DB |||||
QY 5521 AATTCAAGCAAAAGGCGCTCGGGTTGTGCAAAAGGCGCAACGAGGAGGTGCTG 5580
DB |||||
QY 5581 CTCCCGTGTGAGTCCAAAGTGGAGCCCTTGAGACCTTCTGGGCGAGACATGTGGA 5640
DB |||||
QY 5581 CTCCCGTGTGAGTCCAAAGTGGAGCCCTTGGAGACCTTCTGGGCGAGACATGTGGA 5640
DB |||||
QY 5641 ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGCGGAAACCCCGCA 5700
DB |||||
QY 5641 ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGCGGAAACCCCGCA 5700
DB |||||
QY 5701 TAGCATCATTTGATGGATTTACGCTTCTATCATCTAGCCCGCTTACCAACCCAAACACCC 5760
DB |||||
QY 5701 TAGCATCATTTGATGGATTTACGCTTCTATCATCTAGCCCGCTTACCAACCCAAACACCC 5760
DB |||||

QY 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGTGCCCAACTCGCTCTCCACGCGCTGGT 5820
DB |||||
QY 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGTGCCCAACTCGCTCTCCACGCGCTGGT 5820
DB |||||
QY 5821 CAGCTTTCGTGGGCGCGGCATTCGCGGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG 5880
DB |||||
QY 5821 CAGCTTTCGTGGGCGCGGCATTCGCGGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG 5880
DB |||||
QY 5881 TGCTCGTGGACATCTTGGCGGCTATGGGGCAGGGGTAGCGGCGCATCTCGTGGCCTTTA 5940
DB |||||
QY 5881 TGCTCGTGGACATCTTGGCGGCTATGGGGCAGGGGTAGCGGCGCATCTCGTGGCCTTTA 5940
DB |||||
QY 5941 AGGTATGAGCGGCGAGGTGCCCTCCACCGAGGACCTCGTCAACTTCTCCCTGCGCATCC 6000
DB |||||
QY 5941 AGGTATGAGCGGCGAGGTGCCCTCCACCGAGGACCTCGTCAACTTCTCCCTGCGCATCC 6000
DB |||||
QY 6001 TCTCTCTGTGGTCCCTGTGTCGGGCTGTCGGGAGCAATACTGTCGTGCGCACGTGG 6060
DB |||||
QY 6001 TCTCTCTGTGGTCCCTGTGTCGGGCTGTCGGGAGCAATACTGTCGTGCGCACGTGG 6060
DB |||||
QY 6061 GCCCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTCCGTCGCGGGTA 6120
DB |||||
QY 6061 GCCCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTCCGTCGCGGGTA 6120
DB |||||
QY 6121 ACCACGTCTCCCTTACGCACTATGTGCTGAGAGCGACGCTGCAGCAGTGTCACTCAGA 6180
DB |||||
QY 6121 ACCACGTCTCCCTTACGCACTATGTGCTGAGAGCGACGCTGCAGCAGTGTCACTCAGA 6180
DB |||||
QY 6181 TCCTCTTAGCTTACCATCACTCAACTGTGTAAGCGGCTCAACGAGTGGATTAATAGG 6240
DB |||||
QY 6181 TCCTCTTAGCTTACCATCACTCAACTGTGTAAGCGGCTCAACGAGTGGATTAATAGG 6240
DB |||||
QY 6241 ACTGCTTACGCCATGTCTCGGCTGTGGCTAAGGATGTTTGGATTTGGATATGACCG 6300
DB |||||
QY 6241 ACTGCTTACGCCATGTCTCGGCTGTGGCTAAGGATGTTTGGATTTGGATATGACCG 6300
DB |||||
QY 6301 TGTGACTGACTTCAAGACCTGGCTCCAGTCCAAACTCTGCGCGGTTTACCGGAGTCC 6360
DB |||||
QY 6301 TGTGACTGACTTCAAGACCTGGCTCCAGTCCAAACTCTGCGCGGTTTACCGGAGTCC 6360
DB |||||
QY 6361 CTTTCTCTCATGCAACCGCGGTA CAAGGAGTCTGCGGGGGAACGGCATCATGCAAA 6420
DB |||||
QY 6361 CTTTCTCTCATGCAACCGCGGTA CAAGGAGTCTGCGGGGGAACGGCATCATGCAAA 6420
DB |||||
QY 6421 CCACCTGCCATGCGGACAGATCGCGGACATGTCAAAAACGTTCCATGAGGATCG 6480
DB |||||
QY 6421 CCACCTGCCATGCGGACAGATCGCGGACATGTCAAAAACGTTCCATGAGGATCG 6480
DB |||||
QY 6481 TAGGGCTTAGAACCTTGACAGCAACAGTGGCACGGAACGTTCCCATCAACGATACACCA 6540
DB |||||
QY 6481 TAGGGCTTAGAACCTTGACAGCAACAGTGGCACGGAACGTTCCCATCAACGATACACCA 6540
DB |||||
QY 6541 CGGACCTTGCACACCTTCCCGCGGCCAACTATTCAGGGCGCTATGCGGGTGGCTG 6600
DB |||||
QY 6541 CGGACCTTGCACACCTTCCCGCGGCCAACTATTCAGGGCGCTATGCGGGTGGCTG 6600
DB |||||
QY 6601 CTGAGGAGTACGTGAGGTTACGCGTGGGGATTTCCACTACCTGACGGGCATGACCA 6660
DB |||||
QY 6601 CTGAGGAGTACGTGAGGTTACGCGTGGGGATTTCCACTACCTGACGGGCATGACCA 6660
DB |||||
QY 6661 CTGACAAAGTAAAGTCCCATGCGGCTTCCGGCCCCCGAAATTTTTCAGGAGGTGGATG 6720
DB |||||
QY 6661 CTGACAAAGTAAAGTCCCATGCGGCTTCCGGCCCCCGAAATTTTTCAGGAGGTGGATG 6720
DB |||||
QY 6721 GAGTCCGTGTCACAGGTACGCTCCGGCGTCAAAACCTTCTTACGGGAGGACGTACCT 6780
DB |||||
QY 6721 GAGTCCGTGTCACAGGTACGCTCCGGCGTCAAAACCTTCTTACGGGAGGACGTACCT 6780
DB |||||
QY 6781 TCCAGTCCGGCTCAACCAATCTTGGTCCGGTCCGAGCTCCCATCCGAGCCGACCCG 6840
DB |||||
QY 6781 TCCAGTCCGGCTCAACCAATCTTGGTCCGGTCCGAGCTCCCATCCGAGCCGACCCG 6840
DB |||||
QY 6841 ACGTAAACAGTGCTTACTTCCATGCTCACCGATCCCTCCCACTTACAGCAGACGGCTA 6900
DB |||||

6841	Db	 ACGTAACAGTGTCTACTTCATGTCTCAGCATCCCTCCCAACATTACAGCAGAGACGGCTA	6900
6901	Qy	 AGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTTCT	6960
6901	Db	 AGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTTCT	6960
6961	Qy	 CTCGCGCTTTTGAAGGGACATGCACTACCCACCATGACTCCCGGACGCTGACCTCA	7020
6961	Db	 CTCGCGCTTTTGAAGGGACATGCACTACCCACCATGACTCCCGGACGCTGACCTCA	7020
7021	Qy	 TCGAGGCCAACCTTTGTGGCGCAGGAGATGGCGGAAACATCATCTCGGTGGAGTCAG	7080
7021	Db	 TCGAGGCCAACCTTTGTGGCGCAGGAGATGGCGGAAACATCATCTCGGTGGAGTCAG	7080
7081	Qy	 AGAAATAGGTAGTAAATTTCTGCACTCTTTGCAACCGCTTACGGCGAGGGGGATGAGAGGG	7140
7081	Db	 AGAAATAGGTAGTAAATTTCTGCACTCTTTGCAACCGCTTACGGCGAGGGGGATGAGAGGG	7140
7141	Qy	 AGATATCCGTTCGGCGGGAGATCCTCGAAAAATCCAGGAAGTTCCCTCAGCGTTCGCCA	7200
7141	Db	 AGATATCCGTTCGGCGGGAGATCCTCGAAAAATCCAGGAAGTTCCCTCAGCGTTCGCCA	7200
7201	Qy	 TATGGGCACGCCCGGACTACAATCTCTCACTGTCTAGAGTCTCTGAAAGGACCCGGACTACG	7260
7201	Db	 TATGGGCACGCCCGGACTACAATCTCTCACTGTCTAGAGTCTCTGAAAGGACCCGGACTACG	7260
7261	Qy	 TCCTCCGGTGGTACACGGATGCCATTTGCCACTTACCAGGCTCTCTCCAAATACCACTC	7320
7261	Db	 TCCTCCGGTGGTACACGGATGCCATTTGCCACTTACCAGGCTCTCTCCAAATACCACTC	7320
7321	Qy	 CAGCGAAAAAGAGGACGGTTGTCTGTACAGAAATCCAAATGTGTCTTCTGCTTGGCGGAGC	7380
7321	Db	 CAGCGAAAAAGAGGACGGTTGTCTGTACAGAAATCCAAATGTGTCTTCTGCTTGGCGGAGC	7380
7381	Qy	 TCGCCACTAAGACTTCGGTAGTTCGGATTCGTGCGCGTTGTATAGCGGCACCGGACCG	7440
7381	Db	 TCGCCACTAAGACTTCGGTAGTTCGGATTCGTGCGCGTTGTATAGCGGCACCGGACCG	7440
7441	Qy	 CCCTTCGTACCTGGCCCTCCGACGACGGTGACAAAGATCCGAGCTTGAGCTCTACTCCT	7500
7441	Db	 CCCTTCGTACCTGGCCCTCCGACGACGGTGACAAAGATCCGAGCTTGAGCTCTACTCCT	7500
7501	Qy	 CCATGCCCCCTTTGAAGGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGTCCTA	7560
7501	Db	 CCATGCCCCCTTTGAAGGGGAGCGGGGAGCCCGGATCTCAGCGACGGGTCTTGTCCTA	7560
7561	Qy	 CCGTGAGTGAGGAGCTAGTGAGGATTCGTCTGCTCTCTCAATGTCCTATACGTGGACAG	7620
7561	Db	 CCGTGAGTGAGGAGCTAGTGAGGATTCGTCTGCTCTCTCAATGTCCTATACGTGGACAG	7620
7621	Qy	 GCGCCCTGATCAGCCATGCGCTCGGAGGNAAGTACGTGCCCATCAACCCGTTGAGCA	7680
7621	Db	 GCGCCCTGATCAGCCATGCGCTCGGAGGNAAGTACGTGCCCATCAACCCGTTGAGCA	7680
7681	Qy	 ACTCTTTGCTCGCTCACACAAATGTTACTACGCCACAAATCCCGCAGCGCAAGCCTCC	7740
7681	Db	 ACTCTTTGCTCGCTCACACAAATGTTACTACGCCACAAATCCCGCAGCGCAAGCCTCC	7740
7741	Qy	 GGCAGAAAGGTCACTTTTGACAGATTGCAAGTCTCTGGATGATCATTTACGGGACGTAC	7800
7741	Db	 GGCAGAAAGGTCACTTTTGACAGATTGCAAGTCTCTGGATGATCATTTACGGGACGTAC	7800
7801	Qy	 TCAAGGAGATGAAGCGGAGGGTCCACAGTTAAGGCTTAAGCTCTATCTATAGAGAGG	7860
7801	Db	 TCAAGGAGATGAAGCGGAGGGTCCACAGTTAAGGCTTAAGCTCTATCTATAGAGAGG	7860
7861	Qy	 CCTGCAAGCTACGCCCCCAATTTCGGCCAAATCCAAATTTGGCTATGGGGCAAGGACG	7920
7861	Db	 CCTGCAAGCTACGCCCCCAATTTCGGCCAAATCCAAATTTGGCTATGGGGCAAGGACG	7920
7921	Qy	 TCCGGAACCTTATCCAGCAGGGCGGTAAACCATCATCCGCTCCGCTGTGGGAGGACTTGCTGG	7980

Db	7921	TCCGGAACTTATCCAGCAGGGCGGTTAAACCAATCCGCTCCGTGTGGGAGGACTTGTCTGG	7980
Qy	7981	AAGACACTGAAACACCAATTGACACCAACCATCATGSCAAAAAGTGTAGGTTTTCGCGTCC	8040
Db	7981	AAGACACTGAAACACCAATTGACACCAACCATCATGSCAAAAAGTGTAGGTTTTCGCGTCC	8040
Qy	8041	AACCAGAGAAGGAGGCCCAAGCAGCTCGCCTTATCGTATTTCCACAGACCTGGGAGTTTC	8100
Db	8041	AACCAGAGAAGGAGGCCCAAGCAGCTCGCCTTATCGTATTTCCACAGACCTGGGAGTTTC	8100
Qy	8101	GTGTATGCGAGAAGATGGCGCCTTTTACGAGCTGTGTCTCCACCCTTCTTCAGGCCGTGATGG	8160
Db	8101	GTGTATGCGAGAAGATGGCGCCTTTTACGAGCTGTGTCTCCACCCTTCTTCAGGCCGTGATGG	8160
Qy	8161	GCTCTCTATACGGATTTCAATACTCCCTCCCAAGCAGCGGGTTCGATGTTCTTGCGTAATACCT	8220
Db	8161	GCTCTCTATACGGATTTCAATACTCCCTCCCAAGCAGCGGGTTCGATGTTCTTGCGTAATACCT	8220
Qy	8221	GGAATTCAAAGAAATCGCCTATGGGCTTCTCATATGACACCCGCTGTTTTGACTCAACGG	8280
Db	8221	GGAATTCAAAGAAATCGCCTATGGGCTTCTCATATGACACCCGCTGTTTTGACTCAACGG	8280
Qy	8281	TCACTGAGAGTGACATTCGTGTTTGAGGAGTCAATTTACCAATGTGTGACTTGTGCCCCCG	8340
Db	8281	TCACTGAGAGTGACATTCGTGTTTGAGGAGTCAATTTACCAATGTGTGACTTGTGCCCCCG	8340
Qy	8341	AGSCCAGACAGGCCATAAGGTCGCTCACAGCGGCTTTACATCGGGGGTCCCTGACTA	8400
Db	8341	AGSCCAGACAGGCCATAAGGTCGCTCACAGCGGCTTTACATCGGGGGTCCCTGACTA	8400
Qy	8401	ACTCAAAGGGCAGAACTGCGGTTATCGCGGTTCGCGCAAGTGGCGTCTGACGACTA	8460
Db	8401	ACTCAAAGGGCAGAACTGCGGTTATCGCGGTTCGCGCAAGTGGCGTCTGACGACTA	8460
Qy	8461	GCTGCGGTAAATACCTTCACATGTTTACTTTGAAGGCCACTGCAGCGCTGTGAGCTGCAAAAGC	8520
Db	8461	GCTGCGGTAAATACCTTCACATGTTTACTTTGAAGGCCACTGCAGCGCTGTGAGCTGCAAAAGC	8520
Qy	8521	TCCAGGACTGCAGGATGCTCGTGAACGGAGACGACCTTGTGTTATCTGTGAAGCGCGG	8580
Db	8521	TCCAGGACTGCAGGATGCTCGTGAACGGAGACGACCTTGTGTTATCTGTGAAGCGCGG	8580
Qy	8581	GAACCCAGGAGGATGCGGGCGCCTACGAGCCTTACGGAGGCTATGACTAGGTATTTCCG	8640
Db	8581	GAACCCAGGAGGATGCGGGCGCCTACGAGCCTTACGGAGGCTATGACTAGGTATTTCCG	8640
Qy	8641	CCCCCCCCGGGATTCGCCCCCAACCAAGATACGACCTTGAACATCATGTTCTCT	8700
Db	8641	CCCCCCCCGGGATTCGCCCCCAACCAAGATACGACCTTGAACATCATGTTCTCT	8700
Qy	8701	CCAATGTGTGAGTGCAGCGCAGATGCTCTGGCAAAAGGGTATACCTCACCCGTGACC	8760
Db	8701	CCAATGTGTGAGTGCAGCGCAGATGCTCTGGCAAAAGGGTATACCTCACCCGTGACC	8760
Qy	8761	CCACCACCCCCTTGTGACAGGGCTGCGTGGGAGACAGCTAGACACACTTCCAATCAACTCTT	8820
Db	8761	CCACCACCCCCTTGTGACAGGGCTGCGTGGGAGACAGCTAGACACACTTCCAATCAACTCTT	8820
Qy	8821	GGCTAGGCATATCATGATGTATGCGCCCAACCTTATGGCAAGGATGATTTCTGATGACTC	8880
Db	8821	GGCTAGGCATATCATGATGTATGCGCCCAACCTTATGGCAAGGATGATTTCTGATGACTC	8880
Qy	8881	ACTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTTGAAAAGCCCTGGATTGTGAGTCT	8940
Db	8881	ACTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTTGAAAAGCCCTGGATTGTGAGTCT	8940
Qy	8941	ACGGGGCTTGTACTCCATTGAGCCAATTGACCTACTCTCAGATCATTTGAACGACTCCATG	9000
Db	8941	ACGGGGCTTGTACTCCATTGAGCCAATTGACCTACTCTCAGATCATTTGAACGACTCCATG	9000
Qy	9001	GTCTTAGCGCATTTACACTCAGAGTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060
Db	9001	GTCTTAGCGCATTTACACTCAGAGTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060

[illegible][illegible]

QY 3434 CTACTCCCAACAAACGCGGGGCTACTTTGGTTGCATCATCTAGCCTCACAGGCGCGGA 3493
DB |||||
3422 CTATTTCCCAACAAACGCGGGGCTGCTTGGCTGTATCATCTAGCCTCACAGGTCGGA 3481
QY 3494 CAAAGAACAGGTGCAAGGGAGGTTCAAGTGTGTTCTACCGAAACAAATCTTCTCGGC 3553
DB |||||
3482 CAAAGAACAGGTGCAATGGGAGGTTCAAGTGTGTTCTACCGAAACAAATCTTCTCGGC 3541
QY 3554 GACCTGCATCAACGGGTGCTGCTGACCTGTCTACCATGCGCTGCTCGAGACCTTAGC 3613
DB |||||
3542 GACCTGCGTCAATGGCGTGTGTGACCGCTGTACCATGCTGCGGCTCGAAGCCCTGCG 3601
QY 3614 CCGTCCAAAGAGTCCAAATCACCCAAATGTACACCAATGTAGACCAAGGACCTCGTCGGCTG 3673
DB |||||
3602 CGGCCGGAAGGTCCTCAATCACCCAAATGTACACCAATGTAGACCAAGGACCTCGTCGGCTG 3661
QY 3674 GGAGGCGCCCCCGGGGCGCGCTCCATGACACCATGTGAGCTGTGGAGCTCGGACCTTTTA 3733
DB |||||
3662 GCGGCGCCCCCGGGGCGCGCTCCATGACACCGTGCACCTGCGGAGCTCGGACCTTTTA 3721
QY 3734 CTTGTGTACGACATGCTGATGCTCATTCGGGTGCGCGGCGGAGCGACAGAGGGGAAG 3793
DB |||||
3722 CTTGTGTACGAGGCTGTGATGTGCTTCCGGTGC CGCGGCGGCGAGCAGAGGGGGAG 3781
QY 3794 TCTACTCTCCCCAGGCGCGCTCTCTTACCTGAAAGGCTCCTCGGGTGTGTCCTTCTTTG 3853
DB |||||
3782 CTTGCTTTCCCCAGGCGCATCTCTTACCTGAAGGCTCCTCGGGTGGACCACTGCTTTG 3841
QY 3854 CCGTTCGGGGACGTCGTGGGCGCTCTTCGGGCTGCTGTGTGACACCGGGGGGTGCGAA 3913
DB |||||
3842 CCGTTCGGGGACGTTGTAGGCATCTTCGGGCTGCTGTGTGACACCGGGGGGTGCGAA 3901
QY 3914 GCGGTGGAATTCATACCGGTTGAGTCTGTGAACTACCATGCGGTCTCGGCTCTTAC 3973
DB |||||
3902 GCGGTGGAATTCATACCGGTTGAGTCTGTGAAACTACCATGCGGTCTCGGCTCTTAC 3961
QY 3974 AGACAACTCAACCCCCCGGCTGTACCGCAGACATTCCAAAGTGGACATCTGCACGCTCC 4033
DB |||||
3962 AGACAACTCATCCCCCTCGGCGGTACCGCAACATTCCAAAGTGGACATCTGCACGCTCC 4021
QY 4034 TACTGCGAGCGCAAGAGACCAAAAGTCCGCTGCTGTATGAGCCCAAGGTTAAGGT 4093
DB |||||
4022 CACTGCGAGCGCAAGAGACCAAAAGTCCGCTGCTCATGACAGCCCAAGGTTAAGGT 4081
QY 4094 GCTGCTCTGAACCGTTCGCTGCGCCACCTTAGGTTGGGGGTATATCTCAAGC 4153
DB |||||
4082 GCTGCTCTTAAACCGGTCGCTTGC CGCCACATTTGGGCTTTGGAGCGTATATGTCAAGGC 4141
QY 4154 ACACGGTATCGACCCCTAAACATCAGAACTGGGGTAAAGGACCATTTACCGGGGGGCTCCAT 4213
DB |||||
4142 ACATGGCATCGAGCTTAACATCAGAATCGGGTAAAGGACCATCAACCGGGCGGCCCAT 4201
QY 4214 TAGCTACTCCACTATGGCAAGTTCCTGCGCAGGTGCTGTTCGGGGCGCCTATGA 4273
DB |||||
4202 CACGTACTCCACTATTTGAAGTTCCTTGC CAGCGTGGATGCTCGGGGGCGCCTATGA 4261
QY 4274 CATATAATATGTAGTGGTGCCTCAACTGACTGCTACCATCTTTGGGCTCGGCAC 4333
DB |||||
4262 CATATAATATGTAGTGAATGCACTCAACTGACTACCATCTTTGGGCTCGGCAC 4321
QY 4334 AGTCTTGGACCAAGCGGAGACGGCTGGAGCGGGCTGCTGCTGCTCGCCACCGCTACAC 4393
DB |||||
4322 AGTCTTGGATCAGGACAGACGGCTGGAGCGGGCTGCTGCTGCTCGCCACCGCCACGCC 4381
QY 4394 TCCGGGATCGGTTACCGTCCCAACACCCCAATATCGAGGAAATAGGCTCTCAACAATGG 4453
DB |||||
4382 TCCGGGATCGATCACCGTGCACACCCCAACATCGAGGAAGTGGCCTGTCCAACACTGG 4441
QY 4454 AGAGATCCCTTCTATGGCAAGCCATCCCATTTAGGCGCATCAAGGGGGGGAGGCTCT 4513
DB |||||
4442 AGAGATTCCTCTTATGGCAAGCCATCCCATTTAGGCGCATCAAGGGGGGGAGGCTCT 4501

QY 4514 CATTTTCTGCAATTCCAAGAAAGAAATGTGACGAGCTCCCGCAAGAGCTGACAGGCTCGG 4573
DB |||||
4502 CATCTTCTGCCATTCCAAGAAAGAGTGTGACGAGCTCCCGCAAGAGCTGACAGGCTCGG 4561
QY 4574 ACTGAAACGCTGTAGCATATTTACCGGGCTTTGATGTGTGCTCATACCGCTATCGGAGA 4633
DB |||||
4562 ACTCAATGCTGTAGCGTATTTACCGGGGTCTCGATGTGTGCTCATACCGACTAGCGGAGA 4621
QY 4634 CGTCTGTCTGTGGCAACAGACGCTCTAATGACGGGTTTACCGCGCGATTTTGAATCACT 4693
DB |||||
4622 CGTCTGTCTGTGGCAACAGACGCTCTAATGACGGGTTTACCGCGCGATTTTGAATCACT 4681
QY 4694 GATCGACTGCAATATGTGTCAACCGACAGTGTGACTCCAGGAGAACCGCCCTCGGGCAT 4753
DB |||||
4682 GATCGACTGCAACATGTGTCAACCGACAGTGTGATTTAGCTTTGGATCCACCTTCAC 4741
QY 4754 CATTTGAGACGACGACCGTGCCTGAGAGCGCGGTGTGCGCTGCGCAACCGCGAGGTAGAC 4813
DB |||||
4742 CATTTGAGACGACGACCGTGCCTGAGAGCGCGGTGTGCGCTGCGCAACCGCGAGGTAGAC 4801
QY 4814 TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAACCGCCCTCGGGCAT 4873
DB |||||
4802 TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAACCGCCCTCAGGCAT 4861
QY 4874 GTTCGATTTCTCGTCTGTGTGATGTATGACGCGGGCTGTGCTTGTGTATGAGCTCAC 4933
DB |||||
4862 GTTCGACTCTCTCGTCTGTGTGATGTATGACGAGGCTGCGCTTGTGTATGAGCTCAC 4921
QY 4934 GCCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACCTAATAACACCGAGGTTGCCGCTCTG 4993
DB |||||
4922 GCCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACCTAATAACACCGAGGTTGCCGCTCTG 4981
QY 4994 CCAGGACCATCTGGAGTTCTGGAGAGCGTCTTTCACAGGCTCACCCACATAGATGCCCA 5053
DB |||||
4982 CCAGGACCATCTAGAGTTCTGGAGAGCGTCTTTCACAGGCTCACCCACATAGATGCCCA 5041
QY 5054 CTTCTGTCTCCAGACTAAACAGGCGAGAGACAATTTCTTACCTGGTGGCATATCAAGC 5113
DB |||||
5042 CTTCTGTCTCCAGACTAAACAGGCGAGAGACAATTTCTTACCTGGTGGCATATCAAGC 5101
QY 5114 TACAGTGTGCGCAGGGCTCAAGCTCCACCTCATCTGTTGGGACCAAAATGTGGAAGTCT 5173
DB |||||
5102 CACAGTGTGCGCAGGGCTCAGGCTCCACCTCATCTGTTGGGACCAAAATGTGGAAGTCT 5161
QY 5174 CATAGCGTGAACCTCACTGCAAGGCGCAACCCCTGCTGTATAGCTTAGGAGCGCT 5233
DB |||||
5162 CATAGCGTGAACCCCACTGCTGATGGGCAACGCGCTGCTGTACAGGCTAGGAGCGCT 5221
QY 5234 CCAAAATGAGGTCACTCTCACACCCCATAACTAAATACATCATGCGCATGATGTGCGC 5293
DB |||||
5222 TCAAAATGAGGTCACTCTCACACCCCATAACTAAATACATCATGCGCATGATGTGCGC 5281
QY 5294 TGACCTGAGGTGCTCACTAGCACCTGGGTGCTGTAGGCGAGTCTTTCAGCTTTGCG 5353
DB |||||
5282 TGACCTGAGGTGCTCACTAGCACCTGGGTGCTGTAGGCGAGTCTTTCAGCTTTGCG 5341
QY 5354 CGCATACTGCTGACGACAGGCGAGTGTGCTCAATGTGGGCGAGTCACTCTTCTCGGGAA 5413
DB |||||
5342 CGCGTACTGCTGACGACAGGCGGTGTCTATGTGGGCGAGTCACTTGTTCGGGAG 5401
QY 5414 GCCAGCTGCTGCTTCCGACAGGGAAGTCTCTTACCGAGGATTCGATGAGATGGAAGAGTG 5473
DB |||||
5402 GCCAGCTGTTATTCGACAGGGAAGTCTCTTACCGAGGATTCGATGAGATGGAAGAGTG 5461
QY 5474 TGCCTCAACATTTCTTACATCGACAGGGAATGAGCTGCGCGAGCAATTCAGGCAAA 5533
DB |||||
5462 TGCCTTCAACCTTCTTACATCGACAGGGAATGAGCTGCGCGAGCAATTCAGGCAAA 5521
QY 5534 GSCGCTCGGTTGTTGCAAAACGCGCAACGAGGAGGCTGCTGCTCCCGTGTGGA 5593
DB |||||
5522 GSCGCTCGGTTGTTGCAAAACGCGCAACGAGGAGGCTGCTGCTCCCGTGTGGA 5581
QY 5594 GTCCAAGTGGGAGCGCTTTGAGACCTTCTGGGCGAAGCAATGTGGAATTTTCATCAGCGG 5653

QY 7814 GCGAAGCGCTCCACAGTTAAGCTTAAGCTTCTATCTATAGAGGAGGCTGCAAGCTGAC 7873
DB |||||
QY 7802 GCGAAGCGCTCCACAGTTAAGCTTAGCTTCTATCTATAGAGGAGGCTGCAAGCTGAC 7861
DB |||||
QY 7874 GCGCCACANTCGGCAAAATCCAAATTTGGCTATAGGGGCAAGAGAGCTGCGGAACTATC 7933
DB |||||
QY 7862 GCGCCACATTCGGGCAAAATCCAAATTTGGCTATAGGGGCAAGAGAGCTGCGGAGCTATC 7921
QY 7934 CAGCAGGCGCTTAACCATCGCTCCGCTGCTGGGAGGACTTGTGGAGACACTGAAC 7993
DB |||||
QY 7922 CAGCAGGCGCTTAACCATCGCTCCGCTGCTGGGAGGACTTGTGGAGACACTGAAC 7981
QY 7994 ACCAATTGACACACCATCATGCGCAAAAGTAGGTTTCTCGCTCCAAACACAGAGAGGG 8053
DB |||||
QY 7982 ACCAATTGATACACCATCATGCGCAAAAGTAGGTTTCTCGCTCCAAACACAGAGAGG 8041
QY 8054 AGGCGCAAGCAGCTCGCTTATCGTATTCAGACCTGGGAGTTCGTATGCGAGAA 8113
DB |||||
QY 8042 AGGCGCAAGCAGCTCGCTTATCGTATTCAGACCTGGGAGTTCGTATGCGAGAA 8101
QY 8114 GATGCGCTTTACGAGCTGCTCCACCTTCTCAGGCGGTGATGGCTCCTCATACGG 8173
DB |||||
QY 8102 GATGCGCTTTACGAGCTGCTCCACCTTCTCAGGCGGTGATGGCTCCTCATACGG 8161
QY 8174 ATTTCAATCTCCCAAGCAGCGGTGAGTTCTGTTGTAATACCTGGAAATCAAAGAA 8233
DB |||||
QY 8162 ATTCAGTACTCTCTGGCAGCGGTGAGTTCTGTTGTAATACCTGGAAATCAAAGAA 8221
QY 8234 ATGCGCTATGGCTTCTCATATGACACCGCTGTTTGTGACTCAACGGTCACTGAGAGTGA 8293
DB |||||
QY 8222 ATGCGCTATGGCTTCTCATATGACACCGCTGTTTGTGACTCAACGGTCACTGAGAGTGA 8281
QY 8294 CATTGCTGTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGAGGCCAGACAGGC 8353
DB |||||
QY 8282 CATCGGTACTGAGGAATCAATTTACCAATGTTGTGACTTGGCCCCCGAGGCCAGCGGC 8341
QY 8354 CATAGGTCGCTCACAGCGGCTTTACATCGGGGTCCCTGACTTAACCTCAAAAGGCA 8413
DB |||||
QY 8342 CATAGGTCGCTCACAGCGGCTTTATGTCGGGGTCCCTGACTTAATTCGAAGGGCA 8401
QY 8414 GAACTGCGGTATTCGCGGTGCGCGCAAGTGGCGTGTGAGCTAGCTGCGGTAATAC 8473
DB |||||
QY 8402 GAACTGCGGTATTCGCGGTGCGCGCAAGTGGCGTGTGAGCTAGCTGCGGCAACAC 8461
QY 8474 CTTACATGTTACTTGAAGGCACTGCGACCTGTGAGCTGCAAGCTTCCAGACTGAC 8533
DB |||||
QY 8462 CTTACATGTTACTTGAAGGCACTGCGGCTGTGAGCTGCAAGCTTCCAGACTGAC 8521
QY 8534 GATGCTCGTGAAGGAGAGCTTGTGCTGTTATCTGTAAGCGGGAACCCAGGAGGA 8593
DB |||||
QY 8522 GATGCTCGTGAAGGAGAGAGCTTGTGCTGTTATCTGTAAGCTGCGGGAACCCAGGAGGA 8581
QY 8594 TCGGCGGCTTACGAGGCTTACGAGGCTATGACTAGGTATTCGCGCCCCCGGGGA 8653
DB |||||
QY 8582 TCGGCGGCTTACGAGGCTTACGAGGCTATGACTAGGTATTCGCGCCCCCGGGGA 8641
QY 8654 TCGGCCCCAACAGAAATAGCACTGAGCTGATAACATGATGTTCTCCATGTTGTCAGT 8713
DB |||||
QY 8642 CCGGCCCCAACAGAAATAGCACTGAGCTGATAACATGATGTTCTCCATGTTGTCAGT 8701
QY 8714 CCGGCAAGTATGATCTGGCAAAAGGTATACCTCACCGGTGACCCCAACCCCTCT 8773
DB |||||
QY 8702 CCGGCAAGTATGATCTGGCAAAAGGTATACCTCACCGGTGACCCCAACCCCTCT 8761
QY 8774 TGCACGGCTGCTGGGAGACAGCTAGACACACTCCCAATCAACTTTGGCTAGGCAATAT 8833
DB |||||
QY 8762 CGCAGGCTGCTGGGAGACAGCTAGACACACTCCAGTCACTCTGCTAGGCAATAT 8821
QY 8834 CATATGATGCGGCCACCTATGGCAAGGATGATTTCTGATGACTCACTTTTCTCCAT 8893
DB |||||
QY 8822 CATATGATGCGGCCACCTATGGCGAGGATGATTTCTGATGACTCACTTTCTCTAT 8881

QY 8894 CTTTAGCTCAAGAGCAACTTGAAGAACCCCTGGATGTTGAGATCTACGGGCTTGCTA 8953
DB |||||
QY 8882 CTTTAGCTCAAGAGCAACTTGAAGAACCCCTGGATGTTGAGATCTACGGGCTTGCTA 8941
DB |||||
QY 8954 CTCATTTAGGCACTTTGACCTTACCTCAGATCATTTGAACGACTCCATGTTTAGGCAAT 9013
DB |||||
QY 8942 CTCATTTAGGCACTTTGACCTTACCTCAGATCATTTGAACGACTCCATGTTTAGGCAAT 9001
QY 9014 TACACTCCACAGTTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAACT 9073
DB |||||
QY 9002 TTACTCCACAGTTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCATGCTCAGGAACT 9061
QY 9074 TGGGTACCACTTTCGAACTTCGAGACATCGGGCCAGAAAGTGTCCGCTTAAGCTACT 9133
DB |||||
QY 9062 TGGGTACCGCTTTTCGAGTCTGGAGACATCGGGCCAGAAAGTGTCCGCTTAAGCTACT 9121
QY 9134 GTCCAGGGGGGAGGGCGGCACCTTTGTGGAGATACCTCTTTAACTGGGCACTAAGGAC 9193
DB |||||
QY 9122 GTCCAGGGGGGAGGGCTGCCACTTTCGCGCAAGTACCTCTTCAACTGGGCACTAAGAC 9181
QY 9194 CAAGCTTAACTCACTCCCAATCCCGGCGCGTCCAGCTGGACTTGTGGCTGGTTCGT 9253
DB |||||
QY 9182 CAAGCTTAACTCACTCCCAATCCCGGCGCGTCCAGCTGGACTTGTGGCTGGTTCGT 9241
QY 9254 CGCTGTTTACAGCGGGGAGACATATATCACAGCTGTCTGTCGCCAGCCCGCTGGTT 9313
DB |||||
QY 9242 TGCTGTTTACAGCGGGGAGACATATATCACAGCTGTCTGTCGCCAGCCCGCTGGTT 9301
QY 9314 TCGGTTGTCCTTACTCTTCTTCTGTTAGGGGTAGGCAATTTACCTGCTCCCAACCGATG 9373
DB |||||
QY 9302 CATGTTGTCCTTACTCTTCTTCTGTTAGGGGTAGGCAATTTACCTGCTCCCAACCGATG 9361
QY 9374 AACGGGAGCTTAACCACTCCAGGCTTAAGCAATTTCTGTTTCTTTT 9424
DB |||||
QY 9362 AACGGGAGCTTAACCACTCCAGGCTTAAGCAATTTCTTCTTTT 9412

RESULT 3
US-08-324-977-1
; Sequence 1, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Ieao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McIeland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333..9362
US-08-324-977-1

Query Match 85.2%; Score 8170.4; DB 2; Length 9416;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 8636; Conservative 0; Mismatches 776; Indels 0; Gaps 0;
Qy 14 TGGGGGCGACACTCCACCATGATCACTCCCTGTGAGGAACACTACTGCTTCACGCAGAA 73
Db 5 TGGGGGCGACACTCCACCATGATCACTCCCTGTGAGGAACACTACTGCTTCACGCAGAA 64
Qy 74 AGCGTCTAGCCATCGGCTTAGTATGAGTGTGCTGCAGCTCCAGAGCCCCCCCCCTCCCGGG 133
Db 65 AGCGTCTAGCCATCGGCTTAGTATGAGTGTGCTGCAGCTCCAGAGCCCCCCCCCTCCCGGG 124
Qy 134 AGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCGACGACACCGGGTCC 193
Db 125 AGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCGACGACACCGGGTCC 184
Qy 194 TTTCTTGATCAACCCCGCTCAATGCTGGAGATTTGGGCGTGCCTCCCGGAGACTGCTAG 253
Db 185 TTTCTTGATCAACCCCGCTCAATGCTGGAGATTTGGGCGTGCCTCCCGGAGACTGCTAG 244
Qy 254 CCGAGTAGTGTGGTTCGGAAGGCCCTTGTTGTTACTGCTGATAGGTTGCTTGGAGTG 313
Db 245 CCGAGTAGTGTGGTTCGGAAGGCCCTTGTTGTTACTGCTGATAGGTTGCTTGGAGTG 304
Qy 314 CCGGGGAGGTCTCGTAGACCGTGCACCATGACGACGAACTCTAAACCTCAAGAAAC 373
Db 305 CCGGGGAGGTCTCGTAGACCGTGCACCATGACGACGAACTCTAAACCTCAAGAAAC 364
Qy 374 CAACGTAACCAACCCCGCCACAGGACGTCAAGTTCCCGGCGGTGCTCAGATCGT 433
Db 365 CAACGTAACCAACCCCGCCACAGGACGTCAAGTTCCCGGCGGTGCTCAGATCGT 424
Qy 434 TGGTGGAGTTTACCTGTTCCCGCGAGGCCCCAGGTTGGGTGTCGGCGGACTAGGAA 493
Db 425 TGGTGGAGTTTACCTGTTCCCGCGAGGCCCCAGGTTGGGTGTCGGCGGCCCCAGGAA 484
Qy 494 GCGTTCCGAGCGGTGCGAACCTCGTGGAGGGGCAACCTATCCCAAAGGCTCGCCGACC 553
Db 485 GACITCCGAGCGGTGCGAACCTCGTGGAGGGGCAACCTATCCCAAAGGCTCGCCGACC 544
Qy 554 CGAGGCGAGGGCTGGGCTCAGCCCGGGTACCCCTTGGCCCTCTATGCAATAGGGGCT 613

545 CGAGGCGCAGGACCTGGGCTCAGCCCGGTAACCTTGGCCCTCTATGCAATGAGGGCTT 604
Qy 614 GGGGTGGGAGGATGCTCTGTCAACCCCGCGGCTCCCGGCTAGTGGGGCCCCACGGA 673
Db 605 AGGTGGGCGAGGATGCTCTGTCAACCCCGCGGCTCCCGGCTAGTGGGGCCCCACGGA 664
Qy 674 CCCCCGCGTAGGTGCGGTAACCTTGGGTAAAGGTCAATGATACCCCTTACATGCGGCTTCGC 733
Db 665 CCCCCGCGTAGGTGCGGTAACCTTGGGTAAAGGTCAATGATACCCCTTACATGCGGCTTCGC 724
Qy 734 CGATCTCATGGGTACATTCGCTCGTGGGCGCCCCCTTAGGGGCGCTGCGCAGGGGCTT 793
Db 725 CGATCTCATGGGTACATTCGCTCGTGGGCGCCCCCTTAGGGGCGCTGCGCAGGGGCTT 784
Qy 794 GGCACAGGCTGTCGGGTTCTGGAGACGGCGTGAACATATGCAACAGGGAACCTTGCCTGG 853
Db 785 GGCACATGGTGTCCGGGTTCTGGAGACGGCGTGAACATATGCAACAGGGAACCTTGCCTGG 844
Qy 854 TTGCTCTTTCTCTATCT 913
Db 845 TTGCTCTTTCTCTATCT 904
Qy 914 TTATGAAGTGGCAACGCTGTCCGGGATATACCATGTACGAAACGACTGCTTCAACTCAAG 973
Db 905 TTACGAAGTGCACAACGCTGTCCGGGATATATCATGTACGAAACGACTGCTTCAACGCAAG 964
Qy 974 CATTTGTATGAGGACGGAGCGTGTATCATGCTATCTCCGGGTGCGTCCCTGTGTTCA 1033
Db 965 CATTTGTATGAGGACGGAGCTTGTATCATGCTATCTCCGGGTGCGTCCCTGTGTTTCG 1024
Qy 1034 GGAGGGTAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCCAACGCTCCGCGCCAGGAATGC 1093
Db 1025 GGAAGGCAATCTCTCCCGTGTCTGGGTAGCGCTCACTCCCAACGCTCCGACGCGAGACGT 1084
Qy 1094 CAGCGTCCCCCACTACGACAAATACGACGCCACGTCGACTTCTCTGTTGGGACGGTGTCTT 1153
Db 1085 CACCATCCCCCAACGACGATACGACGCCACGTCGACTTCTCTGTTGGGCGGGTGTCTT 1144
Qy 1154 CTGCTCGCTATGTAGCTGGGGGATCTCTGGGATCTATTTTCTCTCTCTCTCTCTCTCTCT 1213
Db 1145 CTGTTCCGCTATGTAGCTGGGGGACCTCTCGGATCTGTTTCTCTCTCTCTCTCTCTCTCT 1204
Qy 1214 CACCTTCTCGCTCGCGGATGAGACAGTGCAGGAGTGCACCTGCTCAATCTATCCCGG 1273
Db 1205 CACCTTCTCGCTCGCGGATGAGACATTTACAGGACTGTAACCTGCTCAATTTATCCCGG 1264
Qy 1274 CCATGTATCAGGTACACCGCATGGCTTGGGATATGATGATGAACCTGGTCACCTCAACAGC 1333
Db 1265 CCATGTGTGGGTACCGGATGGCTTGGGACATGATGATGAACCTGGTCGCCCAACAGC 1324
Qy 1334 CCTAGTGTGTGCGAGTTGCTCCGGATCCCAAGCTGTGCGTGGACATGTTGGGCGGGGC 1393
Db 1325 CCTAGTGTGTGCGAGTTACTCCGGATCCCAAGCGCTGCTGGACATGTTGGGCGGGGC 1384
Qy 1394 CCACTGGGGAGTCTCGCGGCGCTTGGCTACTATTCATGTTAGGGAACCTGGGCTAAGGT 1453
Db 1385 CCACTGGGGAGTCTCGCGGCGCTTGGCTACTATTCATGTTAGGGAACCTGGGCTAAGGT 1444
Qy 1454 TCTGATTTGGGCGTACTCTTTTGGCGGCTTGGACGGGAGAGCCACACGAGCGGGAGGGT 1513
Db 1445 TCTGATTTGATGCTACTTTTGTGGGTGACGGGGATACCCACGTCGACAGCGGGGGC 1504
Qy 1514 GGGCGGCGACACCACTCCGGGTTACGTCCTCTTTTCTCATCTGGGGGCTCTCAGAAAT 1573
Db 1505 GCAAGCCAAACCAACCAAGGCTCGTGTCCATGTTGCAAGTGGGCGCTCTCAGAAAT 1564
Qy 1574 CCAGCTTGTGAATACCAACGCGACCTGGCACATCAACAGGACTGCCCTTAATTCGAATGA 1633
Db 1565 CCAGCTTATAAACCAATGGGAGTTGGCAATCAACAGGACTGCCCTTAATTCGAATGA 1624
Qy 1634 CTCCCTCCAAACTGGGTTCTTTGGCGGCTGTTTTAGCACACAAAGTTCAACTCGTCCGG 1693
Db 1625 CTCTCTCCAGACTGGGTTCTTTGGCGGCTGTTTCTACACATAGTTTCAACTCGTCCGG 1684

[illegible]

Db 8225 AAACCCCATGGCTTTTTCATATGACACTCGCTGTTTCGACTCAACGGTCAACCGAGAACGA 8284
Qy 8294 CATTCTGTTGAGGAGTCAATTTACCAATGTTGAGCTTTGGCCCCCGAGGCCAGACAGGC 8353
Db 8285 CATCCGTTGAGGAGTCAATTTACCAATGTTGAGCTTTGGCCCCCGAGGCCAGACAGGC 8344
Qy 8354 CATAAGTTCGCTCACAGAGCGGCTTTTACATCGGGGGTCCCTGACTAACTCAAAAGGGCA 8413
Db 8345 CATAAAATCGCTCACAGAGCGGCTTTTATATCGGGGGTCCCTGACTAAATCAAAAGGGCA 8404
Qy 8414 GAATCGGGTTATCGCGGTTCGCGGCAAGTGGCGTCTGACACTAGCTGCGGTATAC 8473
Db 8405 GAATCGGGTTATCGCGGTTCGCGGCGAGCGCGTCTGACACTAGCTGCGGTATAC 8464
Qy 8474 CCTCACATGTTACTTGAAGGCCACTGCGAGCTGTGCGAGCTGCAAAAGCTCCAGGACTGCAC 8533
Db 8465 CCTCACATGTTACTTGAAGGCCCTTCGAGCCTGTGCGAGCTGCGAGCTCCAGGACTGCAC 8524
Qy 8534 GATGCTCGTGAACGGAGACGACCTTGTGTTATCTGTGAAGCGCGGGAACCCAGGAGGA 8593
Db 8525 GATGCTCGTGAACGGAGACGACCTCGTGTGTTATCTGTGAAGCGCGGGAACCCAGGAGGA 8584
Qy 8594 TGGCGGGCCCTTACGAGCCTTACGGAGGCTATGACTAGTATTCGCGCCCCCGCGGGA 8653
Db 8585 CGCGGGCAGCCTACGAGCTTTCAGGAGGCTATGACTAGTATTCGCGCCCCCGCGGGA 8644
Qy 8654 TCCGCCCCCAACCAAGATACGACCTGGAGCTGATAACATCATGTTCTCCCAATGTGTCAGT 8713
Db 8645 CCGCGCCCCCAACCAAGATACGACCTGGAGCTGATAACATCATGTTCTCCCAATGTGTCAGT 8704
Qy 8714 CGCGCAGATGCTATCTGGCAAAAGGTTATCTACTCTCACCGGTGACCCACCCACCCCT 8773
Db 8705 CGCCACGATGCTATCAGGCAAAAGGTTATCTACTCTCACCGGTGATCCACCCACCCCT 8764
Qy 8774 TGCAACGGGCTCGTGGGAGACAGCTAGACACACTCCCAATCACTCTTGGCTAGGCAATAT 8833
Db 8765 AGCACGGGCTCGTGGGAGACAGCTAGACACACTCCCAATCACTCTTGGCTAGGCAATAT 8824
Qy 8834 CATCATGTATGCGGCCACCTTATGGCAAGATGATCTGATGACTCACTTTTCTCCAT 8893
Db 8825 TATTATGTATGCGGCCACCTTGTGGCAAGATGATCTGATGACTCACTTTTCTCCAT 8884
Qy 8894 CTTTCTAGCTCAAGACCACTTTGAAAAGCCCTGGATTGTGAGATCTACGGGGCTTGCTA 8953
Db 8885 CCTTCTAGCGGAGGACACTTTGAAAAGCCCTGGATTGTGAGATCTACGGGGCTTGCTA 8944
Qy 8954 CTCCATTGAGCCACTTGACCTACCTCAGATCAATTTGAAGCACTCCATGCTTTAGCGCAT 9013
Db 8945 CTCCATTGAGCCACTTGACCTACCTCAGATCAATTTGAAGCACTCCATGCTTTAGCGCAT 9004
Qy 9014 TACATCTCACAGTTACTCTCCAGGTGAGATCAATAGGTGCTTCATGCTCAGGAACT 9073
Db 9005 TTCACTTCAATGTTACTCTCCAGGTGAGATCAATAGGTGCTTCATGCTCAGGAACT 9064
Qy 9074 TGGGGTACCCCTTGGCAACTGAGACATCGGSCCAGAGTGTCCGGCTGAAGTACT 9133
Db 9065 TGGGGTACCCCTTGGCAACTGAGACATCGGSCCAGAGTGTCCGGCTGAAGTACT 9124
Qy 9134 GTCCAGGGGGGAGGGCGGCACCTTGTGGCAGATACCTTTTAACTGGGCGAGTAAGGAC 9193
Db 9125 GTCCAGGGGGGAGGGCGGCACCTTGTGGCAGATACCTTTTAACTGGGCGAGTAAGGAC 9184
Qy 9194 CAAAGTTAACTCACTCCAAATCCCGCGCGGTCCAGCTGAGTGTCTGGCTGGTTGCT 9253
Db 9185 CAAACTTAACTCACTCCAAATCCCGCGCTGCGTCCCGGTGGACTGTCCGGTGGTTGCT 9244
Qy 9254 CGCTGTTTACAGCGGGGAGACATATATCAGGCTGTCTCGTCCCGACCCCGCTGTT 9313
Db 9245 TGCTGGTTTACAGCGGGGAGACATATATCAGGCTGTCTCGTCCCGACCCCGCTGTT 9304
Qy 9314 TCCGTTGTGCTTACTCTTCTGTTAGGGGTAGGCAATTTACCTGCTCCCAACCGCATG 9373

Db 9305 CATGCTGTGCTTACTCTTCTGTTAGGGGTAGGCATCTACCTGCTCCCAACCGATG 9364
Qy 9374 AACGGGAGCTAAACCACTCCAGGCTTAAGCCATTTCTGTTTTTTTTTTTTT 9425
Db 9365 AACGGGAGATAAACACTCCAGGCAATAGGCACTCCCGCTTTTTTTTTTTTTT 9416
RESULT 4
US-08-384-616-1
; Sequence 1, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 333..9362
US-08-384-616-1
Query Match 85.2%; Score 8170.4; DB 2; Length 9416;
Best Local Similarity 91.8%; Pred. No. 0;

Matches 8636; Conservative 0; Mismatches 776; Indels 0; Gaps 0;			
QY	14	TGGGGCGGACACTCACACATGAATCACTCCCTGTGAGGAATCTACTGTCTTCAACGAGAA	73
Db	5	TGGGGGCGACACTCACACATAGATCACTCCCTGTGAGGAATCTACTGTCTTCAACGAGAA	64
QY	74	AGCGTCTAGCCATGGCGTGTAGTATGAGTGTCTGTGAGCCTCCAGGACCCCCCTCCCGGG	133
Db	65	AGCGTCTAGCCATGGCGTGTAGTATGAGTGTCTGTGAGCCTCCAGGACCCCCCTCCCGGG	124
QY	134	AGAGCCATAGTGGTCTGCGGAAACCGGTGAGTATACACCGGAATTTGCCAGACGACGGGTCC	193
Db	125	AGAGCCATAGTGGTCTGCGGAAACCGGTGAGTATACACCGGAATTTGCCAGACGACGGGTCC	184
QY	194	TTTTCTGGATCAACCGGCTCAATGCTGGAGATTTTGGGGTCCCGCCGAGACTGCTAG	253
Db	185	TTTTCTGGATCAACCGGCTCAATGCTGGAGATTTTGGGGTCCCGCCGAGACTGCTAG	244
QY	254	CCGAGTAGTGTGGGTGCGGAAAGGCTTGTGGTACTGTGCTGATAGGGTGTCTTGGAGTG	313
Db	245	CCGAGTAGTGTGGGTGCGGAAAGGCTTGTGGTACTGTGCTGATAGGGTGTCTTGGAGTG	304
QY	314	CCCCGGGAGTCTCGTAGACCGTGACACCATGAGCAAGAACTCTAAACCTCAAAGAAAAAC	373
Db	305	CCCCGGGAGTCTCGTAGACCGTGACACCATGAGCAAGAACTCTAAACCTCAAAGAAAAAC	364
QY	374	CAACGTAAACCAACCGCCGCCACAGACGCTCAAGTTCCTGGGGGGTGTTCAGATCGT	433
Db	365	CAACGTAAACCAACCGCCGCCACAGACGCTCAAGTTCCTGGGGGGTGTTCAGATCGT	424
QY	434	TGGTGGAGTTTACCTGTTCCTCGCGAGGGGCCACAGGTTGGGTGTGCGCGCGACTAGGAA	493
Db	425	TGGTGGAGTTTACCTGTTCCTCGCGAGGGGCCACAGGTTGGGTGTGCGCGCGCCAGGAA	484
QY	494	GGCTTCCGAGCGGTGCGAACTCGTGGAAAGGCGAACACTATCCAAAGGCTGCGCGACC	553
Db	485	GACTTCCGAGCGGTGCGAACTCGTGGAAAGGCGAACACTATCCCAAGGCTGCGCGGCC	544
QY	554	CGAGGCGAGGCTGGGCTCAGCCCGGGTACCTTTGGCCCCCTCTATGGCAATGAGGGCT	613
Db	545	CGAGGCGAGGCTGGGCTCAGCCCGGGTACCTTTGGCCCCCTCTATGGCAATGAGGGCTT	604
QY	614	GGGTGGGCGAGATGGCTCTCTGTCAACCGCGGCTCCCGGCTCTAGTTGGGGCCCGACGGA	673
Db	605	AGGTGGGCGAGATGGCTCTCTGTCAACCGCGGCTCCCGGCTCTAGTTGGGGCCCGACGGA	664
QY	674	CCCCGGCGTAGTTCGCGTAACTTGGGTAAAGTTCATCGATACCCCTTACATGCGGGTTCGC	733
Db	665	CCCCGGCGTAGTTCGCGTAACTTGGGTAAAGTTCATCGATACCCCTCACAATGCGGCTTCGC	724
QY	734	CGATCTCATGGGGTACATTCGCGTGTGCGGCGCCCCCTTAGGGGGGCTGCGAGGGCCTT	793
Db	725	CGATCTCATGGGGTACATTCGCGTGTGCGGCGCCCCCTTAGGGGGGCGCTGCGAGGGCCTT	784
QY	794	GGCACACGGTGTCCGGGTTCTGGAGACGGCGTGAACATATGCAACAGGGGAACTTCCCGGG	853
Db	785	GGCACATGGTGTCCGGGTTCTGGAGACGGCGTGAACATATGCAACAGGGAACTTCCCGGG	844
QY	854	TTGCTCTTTCTCTATCTCCTCTTGGCTCTGTGCTCTGTGTTTGAACATCCAGCTTCGCG	913
Db	845	TTGCTCTTTTCTATCTTCTCTTGGCTCTGTGCTCTGTGTTTGAACATCCAGCTTCGCG	904
QY	914	TTATGAGTGCACAGCTTCGGGATATACATGTACAGAACGACTGCTCAACTCAAG	973
Db	905	TTACGAAGTGCAACAGCTTCGGGATATATCATGTACAGAACGACTGCTCAACGCGAAG	964
QY	974	CATTGTGATAGGACGCGAGCTGATCATGATACTCCCGGGTGGTGCCTGTGTGTCA	1033
Db	965	CATTGTGATAGGACGCGAGCTTGTATCATGATACTCTGGGTGGTGCCTGTGTGTGG	1024
QY	1034	GGAGGGTAAACAGCTCCCGTGTGGGTAGCGCTCACTCCACGCTCGCGGCCAGGAATGC	1093
Db	1025	GGAAAGCAACTCTCTCCCGTGTGGGTAGCGCTCACTCCACGCTCGCAGGCCAGGAACGT	1084

QY	1094	CAGCGTCCCCACTACGACAAATACGACGCGACGTGCGACTTGTCTGTTGGGAGCGGCTGCTTT	1153
Db	1085	CACATATCCCAACACGACGATACGACGCCACGTGATCTGCTGTTGGGGGGGCTGCTTT	1144
QY	1154	CTGCTCGCTATGTACGTGGGGGATCTCTGCGGATCTATTTTCTCTGCTCTCCAGCTGTT	1213
Db	1145	CTGTTCCGCTATGTACGTGGGGGACCTCTGCGGATCTGTTTTTCTCTCTCTCAGCTGTT	1204
QY	1214	CACCTTCTCGGCTCGCGGCAATGAGACGTGAGGACTGCAACTGCTCAATCTATCCCGG	1273
Db	1205	CACCTTCTCGCTCGCGGCAATGTACATTAACGAGCTGTAACTGCTCAATTTATTCGCGG	1264
QY	1274	CCATGTATCAGGTCAACGCAATGGCTTGGGATATGATGATGAACCTGCTCACCTACAAACAGC	1333
Db	1265	CCATGTGTGGGTCAACGCTTGGGATGATGATGAACCTGCTGCCCCACAAACAGC	1324
QY	1334	CCTAGTGGTGTGCGAGTTGCTCCGATCCCAAGCTGTCTGGACATGCTGGCGGGGCG	1393
Db	1325	CCTAGTGGTGTGCGAGTTTACTCCGATCCCAAGCCGTCTGGACATGCTGGCGGGGCG	1384
QY	1394	CACTGGGGAGTCTTGGCGGCTTGGCTACTATTTCCATGGTAGGGAACCTGGGCTAAAGT	1453
Db	1385	CACTGGGGAGTCTTGGCGGCTTGGCTACTATTTCCAATGGCGGGGAACTGGGCTAAAGT	1444
QY	1454	TCTGATTTGTGGCGCTACTCTTTTGGCGGTTGACGGGAGACCCACACAGCGGGAGGCT	1513
Db	1445	TCTGATTTGTGATGTACTTTTTTGGCGTTGACGGGATACCCACGTGACAGGGGGGCG	1504
QY	1514	GGCGGCGCACACCTCCGGGTTCACTCCCTTTTCTCATCTGGGGGCTCTCGAAAT	1573
Db	1505	GCAAGCCAAACACCAACAGCTCGTGTCTCATGTTCCCAAGTGGGCGCTCTCGAAAT	1564
QY	1574	CCAGTCTGTGAATACCAACGCGAGCTGGCAATCAACAGGACTGCCCCCTAAATTTGCAATGA	1633
Db	1565	CCAGCTTATAAACCAACCAATGGGAGTTGGCACATCAACAGGACTGCCCCCTGAACTGCAATGA	1624
QY	1634	CTCCCTCCAAACTGGGTCTTTTGGCGGCTGTTTACGCACACAGTTTCACTCTGCTCGG	1693
Db	1625	CTCTCTCAGACTGGGTCTTTTGGCGGCTGTTTACACACATAGTTTCAACTCTGCTCGG	1684
QY	1694	GTGCCGGAGCGCATGGCCAGCTGCGGCCCATTTGACTGGTTTCCGCCAGGGGTGGGGGCC	1753
Db	1685	GTGCCAGAGCGCATGGCCAGCTGCGGCCCATTTGACAAGTTTCGACAGGGATGGGTCC	1744
QY	1754	CATCACCTATATTAAGCCTAAACAGCTCGGATCAGAGGCTCTTATTTGCTGGCATTAACGCGC	1813
Db	1745	CATTACTTATGCTGAGTCTAGCAGATCAGACAGAGGCCATATTGCTGGCACTACCCACCC	1804
QY	1814	TCGACCGTGTGGTGTGCTACCGCGTGTGAGTGTGTTCCAGTGTATTTTTCACCCG	1873
Db	1805	TCCCAATTTGATACCTGCTGCTGAGGAGTGTGCGGGCCGAGTGTCTGCTTCAACCCG	1864
QY	1874	AAGCCCTGTGTGTGGGGGACACCGATCTGTTTCCGGTGTCCCTACGCTATAGCTGGGGGA	1933
Db	1865	AAGCCCTGTGTGTGGGGGACACCGATCTGTTTCCGGTGTCCCTACGCTATAGATGGGGGA	1924
QY	1934	GAATGAGACAGAGTGTGCTCTCAACAAACGCGGTCCGCCCAAGGCAACTGTTTGG	1993
Db	1925	GAAACGAGACTGACGCTGCTGCTCAACAAACGCGGCGCGCAAGGCAACTGTTTGG	1984
QY	1994	CTGTACATGGATGAATAGTACTGGGTTCACTAAGACGTGCGGAGGTCCCGCTGTAAACAT	2053
Db	1985	CTGCACATGGATGAATAGCACCAGGTTTCAACAAGACATGTGGGGGGCCCCCGTGTAAACAT	2044
QY	2054	CGGGGGGCTCGGTAAACCGACCTTGATCTGCCCAACGAGCTGCTTCCGGAAGCACCCCGA	2113
Db	2045	CGGGGGGCTCGGCAACAACAACCTGACCTGCCCAACGAGCTGCTTCCGGAAGCACCCCGA	2104
QY	2114	GGCTACTTACAAATGTGCTCGGGGCCCTGGTTGACACTAGGTGCTTAGTAGACTA	2173
Db	2105	GGCTACTTACAAATGTGTTTGGGGCCCTTGGCTGACACCTAGGTGCTAGTTGACTA	2164

QY 2174 CCCATACAGCCTTTGGCACTACCCCTGCCACTCTCAATTTTCCATCTTTAAAGTTAGGAT 2233
DB 2165 TCCATACAGCCTCTGGCATTTACCCCTGCCACTGTAACTTTTACCATCTTCAAGGTTAGGAT 2224
QY 2234 GTATGTGGGGGGCGTGGAGCACAGGCTCAATGCCGATGCAATTTGAGCTCGAGGAGCG 2293
DB 2225 GTATGTGGGGGGGTGGAGCACAGGCTCAATGTGTGATGCAATTTGACCCGAGGAGCG 2284
QY 2294 CTGTAACTTTGGAGGACAGGAGTATAGTTCAGAACTCAGCCCGCTGCTGTCTGTACACAG 2353
DB 2285 TTGTGACTTGGAGGACAGGAGTATAGTTCAGAACTCAGCCCGCTGCTGTCTGTACACAG 2344
QY 2354 GTGGCAGATACGCTCTGCTGTCTTACCACCTTACCGGCTTTATCCACTGGTTTGTATCCA 2413
DB 2345 GTGGCAGGTAATGCTGCTGTCTTACCACCTTACCGGCTTGTCCACTGGCTGTGATTTCA 2404
QY 2414 TCTCCATCAGAACTCGTGACGTCGAACTACCTGTACGGGTAGGGTACAGGCTTGTCTC 2473
DB 2405 CTTCCATCAGAACTCGTGACGTCGAACTACCTGTACGGGTAGGGTACAGGCTTGTCTC 2464
QY 2474 CTTTGAATCAAAATGGGAGTACATCTCTGTCTTCTCTCTCTGCGCAGACGCGCGT 2533
DB 2465 CTTTGAATCAAAATGGGAGTATGCTCTGTCTTCTCTCTCTAGCGGACGACGCTGT 2524
QY 2534 GTGTGCTGCTTGTGAGTATGCTGTCTGATAGCCAGGCTGAGGCGGCTTACAGAACTT 2593
DB 2525 CTGTGCTGCTTGTGATGATGCTGTCTGATAGCCAGGCGGAGCGGCTTGGAGAACCT 2584
QY 2594 GGTGTCTCAATCGCGCGTCCGTGGCGGAGCGCATGTATCTCTCTCTCTCTGTT 2653
DB 2585 GGTGGTCTCAATCGCGCGTCTGTGGCGGAGCGCATGTATCTCTCTCTCTCTGTT 2644
QY 2654 CTTTCTGCGCGCTGTGTACATTAAGGCGAGGCTGGCTCTCTGGGCGGCGTATGCTTTTA 2713
DB 2645 CTTTCTGCGCGCTGTGTACATTAAGGCGAGGCTGGCTCTCTGGGCGGCGATATGCTTTTA 2704
QY 2714 TGGCGTATGCGCGCTGTCTGTCTCTCTGTGGCGGTTACACAGAGCTTACGCTTGA 2773
DB 2705 TGGCGTATGCGCGCTGTCTGTCTCTGTGGCGGTTACACAGAGCTTACGCTTGA 2764
QY 2774 CCGGAGATGGCTGCATCTGCGGGGTGCGGTTCTTGTAGGCTGCTGATTTCTTGACCTT 2833
DB 2765 CCGGAGATGGCTGCATCTGCGGGGTGCGGTTCTTGTAGGCTGCTGATTTCTTGACCTT 2824
QY 2834 GTACCACTATCAAAAGTGTCTCACTAGGCTCATATGGTGTGTACAATCTTTATCAC 2893
DB 2825 GTACCACTATCAAAAGTGTCTCACTAGGCTCATATGGTGTGTACAATCTTTATCAC 2884
QY 2894 CAGAGCGAGCGGCACATGCAAGTGTGGTCTCCCTCCCTCAACGTTTGGGAGCGCGGA 2953
DB 2885 CAGAGCGAGCGGCACATGCAAGTGTGGTCTCCCTCCCTCAACGTTTGGGAGCGCGGA 2944
QY 2954 TGCCATCATCTCTCTCAGTGTGGGTTTCACTCAGAGTAAATTTTGCATCAACCAACT 3013
DB 2945 TGCCATCATCTCTCTCAGTGTGGGTTTCACTCAGAGTAAATTTTGCATCAACCAACT 3004
QY 3014 CTTGTCGCCATCTCGSCCGCTCATGTGCTCCAGGCTGGCATACAGAGTGCCTGA 3073
DB 3005 TCTAATTTGCCATCTCGTCTCGCTCATGTGCTCCAGTGTGCTCAAGTGTGCTCAAGTGTGCTGA 3064
QY 3074 CTTGTCGCCCTCAAGGCTCATTTGTCGATGATGTTAGTGTGCAAAAGTTCGCGGGG 3133
DB 3065 CTTGTCGCCCTCAAGGCTCATTTGTCGATGATGTTAGTGTGCAAAAGTTCGCGGGG 3124
QY 3134 TCATTAATGCTCAAAATGGTCTTCAAGTGTGGCGGCTGTACAGGTACGTATGTTATTA 3193
DB 3125 TCATTAATGCTCAAAATGGTCTTCAAGTGTGGCGGCTGTACAGGTACGTATGTTATTA 3184
QY 3194 CCATCTTACCCACTGCGGGACTGGCCCAACGCGGCTTACGACCTTGGGTGGCGGT 3253
DB 3185 CCATCTTACCCCGTACGCGGATTTGGCCACGCGCGGCTTACGACCTTGGGTGGCGGT 3244
QY 3254 AGAGCCCGTCTCTTCTCCGCCATGGAGACCAAGGTCATCACCTGGGAGCAGACACCGC 3313

DB 3245 GGAGCCCGTCTCTCTCCGACATGGAGACCAAGATCATCACCTGGGAGCAGACACCGC 3304
QY 3314 TCGGTGTGGGACATCATCTTGGGTCTACCCGTCTCCGCCGAGGGGAGAGATATTT 3373
DB 3305 GCGGTGTGGGACATCATCTTGGGTCTGCCCGTCTCCGCCGAGGGGAAAGAGATACT 3364
QY 3374 TTTGGACCGGCTGTAGTCTCGNAGGCAAGGTCGCGACTCTTTGCGCCCATCACGGC 3433
DB 3365 CTTGGGCCCGGCGATAGTCTTGAAGGGCGGGGTTGCGACTCTCGCGCCATCACGGC 3424
QY 3434 CTACTCCCAACAAACCGCGGGCGTACTTGGTTGCTCATCATCTAGCTTCAAGGCCGGGA 3493
DB 3425 CTACTCCCAACAGACGCGGGGCTACTTGGTTGCTCATCATCTAGCTTCAAGGCCGGGA 3484
QY 3494 CAAGAACCAAGTTCGAAGGGAGGTTCAAGTGGTTCTACCGCAACACAATCTTTCTTGGC 3553
DB 3485 CAAGAACCAAGTTCGAAGGGAGGTTCAAGTGGTTCTACCGCAACACAATCTTTCTTGGC 3544
QY 3554 GACCTGCATCAAGGGCTGTGTGGACTGTCTACCATGGCGCTCGAAGACCCCTAGC 3613
DB 3545 GACCTGCATCAAGGGCTGTGTGGACTGTCTACCATGGCGCTCGAAGACCCCTAGC 3604
QY 3614 CGGTCAAAAGGTCCAATCAACCAATGTACACCAATGTAGACCTTGGACCTCGTGGCTG 3673
DB 3605 CGGCCAAAGGGGCAATCACCCAGATGTACATTAATGTGGAACAGGACCTCGTGGCTG 3664
QY 3674 GCAGGCGCCCCCGGGCGGCTCCATGACACCATGTCAGCTGTGGCAGCTTGAACCTTTA 3733
DB 3665 GCCAAAGCCCCCGGGCGGCTTCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTA 3724
QY 3734 CTTTGGTCAAGAGACATGCTGATGTCTTCCGTGCGCGGAGGCGACAGCGGGGAAG 3793
DB 3725 CTTTGGTCAAGAGACATGCTGATGTCTTCCGTGCGCGGAGGCGACAGTATGGGGGAG 3784
QY 3794 TCTACTCTCCCGAGGCTCTCTTACCTGAAAGGCTCTCGGGTGGCTCATTTGCTTTT 3853
DB 3785 CTTGCTCTCCCGAGGCTCTCTTCTTCTTGAAGGCTCTTCCGGTGGTCCACTGCTCTG 3844
QY 3854 CTTTGGGGCAGCTCGTGGGCGCTTCTCGGGTGTGTGTGCAACCGGGGGTTCGGAA 3913
DB 3845 CTTTGGGGCAGCTCGTGGGCGCTTCTCGGGTGTGTGTGCAACCGGGGGTTCGGAA 3904
QY 3914 GCGGTGGATCTTCAATACCGCTTGTGATGTGGAATCTACCATGGGCTCTCGGCTTCAC 3973
DB 3905 GCGGTGGATCTTGTGCGCTGTAGATCCATGGAACTACTATGCGGCTCTCGGCTTCAC 3964
QY 3974 AGACAACTCAACCCCGGCTGTACCGCAGACATTCGAAGTGGCACAATCTGCAAGCTCC 4033
DB 3965 GGAACACTCAATCCCGGCGGCTGTACCGGCTGTACCGGCTCATTTCAAGTGGCCACCTTACAGCTCC 4024
QY 4034 TACTGGCAGCGGCAAGAGCACCAAGTGCAGGCTGTGTATGACGCCCAAGGGTCAAGGT 4093
DB 4025 CACTGGCAGCGGCAAGAGTACTTAAAGTGCAGGCTGTATATGACGCCCAAGGGTCAAGGT 4084
QY 4094 GTCGTCTGAAACCGCTCGCTGTGCGGCACTTATAGGGTTGGGGGCTATATGTCCAAGGC 4153
DB 4085 GTCGTCTCAATCCGTCTCGCTGTGCGCTACCTTATAGGGTTGGGGGCTATATGTCTAAGGC 4144
QY 4154 ACAGGTATCGACCTTAACTCAGAACTGGGGTAAAGACCAATACCAAGGCGGCTCCAT 4213
DB 4145 ACAGGTATCGACCTTAACTCAGAACTGGGGTAAAGACCAATACCAAGGCGGCTCCAT 4204
QY 4214 TACGTACTCAACCTTATGGCAAGTTCCTTGGCGAGCGGTGGCTGTCTGGGGGCGCTATGA 4273
DB 4205 CACATCTTCACTTATGGCAAGTTCCTTGGCGATGGTGGTGTCTGGGGGCGCTATGA 4264
QY 4274 CATCATTAATGTGATGATGCCACTCACTCACTGACTCACTACCATCTTTGGGCACTGGCAC 4333
DB 4265 CATCATTAATGTGATGATGCCACTCACTCACTGACTCACTACCATCTTTGGGCACTGGCAC 4324
QY 4334 AGTCTGTGACCAAGCGGAGGCTGTGAGCGGCTGTCTGTCTGCTCGCCACCGCTTACAC 4393

Db 4325 AGTCTGGACCAAGCGGAGACGGCTGGAGCGCGGCTTGTCTGTCTCGCCACCGCTACGCC 4384
Qy 4394 TCCGGGATCGGTTACCGTGGCCACACCCCAATATCCAGGAAATAGGCTGTCCAAATGG 4453
Db 4385 TCCGGGATCGGTCACCGTGGCCACACCCAAATATCAGGAGGTGGCCCTGTCTAATACCTGG 4444
Qy 4454 AGAGATCCCTTCTATGGCAAGACCATCCCCATTGAGCGCCATCAAGGGGGGAGGCATCT 4513
Db 4445 AGAGATCCCTTCTATGGCAAGACCATCCCCATTGAGCGCCATCAGGGGGGAGGCATCT 4504
Qy 4514 CATTTTCTGCCATCCAAAGAAATGTGAAGAGTTCGCGCAAAAGCTGACAGGCTCGG 4573
Db 4505 CATTTTCTGTCAATCCAAAGAAAGTGCAGCAGCTCGCGAAAGCTGTCAAGGCTCGG 4564
Qy 4574 ACTGAACGCTGTAGCATATTACCGGGGCTTGATGTGTCCGTCATACCGCTTATCGGAGA 4633
Db 4565 AATCAACGCTGTGGCGTATTACCGGGGCTGATGTGTCCGTCATACCACCTATCCGAGA 4624
Qy 4634 CGTCTGTGTGGGCAACAGACGCTTAATGACGGTTCACCGCGATTTTGACTCAGT 4693
Db 4625 CGTCTGTGTGGCAACAGACGCTCTGATGACGGCTATACGGGCGACTTTGACTCAGT 4684
Qy 4694 GATCAGCTGCAATACATGTGTCAACAGACAGTTCAGCTTTGGATCCACACTTCAC 4753
Db 4685 GATCAGCTGTAACATATGTGTCAACAGACAGTTCAGCTTTGGATCCACACTTCAC 4744
Qy 4754 CATTGAGACGACGCGTCCCCAAGACGCGGTGTTCGCGCTCGAACCGCGAGGTAGAAC 4813
Db 4745 CATTGAGACGACGCGTCCCCAAGACGCACTGTTCGCGCTCGAACCGCGGCGGTAGGAC 4804
Qy 4814 TCGCAGGGGTAGAGTGGCATCTACAGGTTTGTGATCCAGGAGACGSCCTCGGGCAT 4873
Db 4805 TGGCAGGGGTAGAGAGGCAATCTACAGTTTGTGACTCCGGAGAACGGCCCTCGGGCAT 4864
Qy 4874 GTTCGATTTCTCGGCTCTGTGTGAGTGTATGACGCGGCTGTGTCTGTGTGATGAGCTCAC 4933
Db 4865 GTTCGATTTCTCGGCTCTGTGTGAGTGTATGACGCGGCTGTGTCTGTGTGATGAGCTCAC 4924
Qy 4934 GCGCGTGTAGACCTCGGTTAGTTGGGCTTTACCTAAATACACAGGTTTCCCGTCTG 4993
Db 4925 CCGGCGGAGACCTCGGTTAGTTGGGCTTACCTGMAACACACAGGTTTCCCGTTTG 4984
Qy 4994 CCAGGACCAATCTGGAGTCTCGGAGAGCGTCTTACAGGCTTACCCACATAGATGCCA 5053
Db 4985 CCAGGACCACTGGAGTCTGGAGAGTGTCTTACAGGCTTACCCATATAGATGCACA 5044
Qy 5054 CTTCTGTCTCCAGACTTAAACAGGACGAGAGCAACTTTCTTACCTGGTGGCATATCAAGC 5113
Db 5045 CTTCTGTCTCCAGACCAAGCAGGACGAGAGCAACTTTCCCTACCTGGTAGCATACCAAGC 5104
Qy 5114 TACAGTGTGGCGAGGCTCAAGCTCCAGTCTGATGTGGGACCAATGTGGAAGTGTCT 5173
Db 5105 CACGCTGTGGCGAGGCTCAGGCCACCTCCATCATGGGATCAAAATGTGGAAGTGTCT 5164
Qy 5174 CATACGGCTGAAACCTTACATCTCACGGGCCAACACCCCTGCTGTATAGGCTAGGAGCCGT 5233
Db 5165 CATACGGCTGAAACCTTACGCTGACGGGCCAACACCCCTGCTGTATAGGCTGGAGCCGT 5224
Qy 5234 CCAAAATAGGCTCATCTCTCACACACCCCATAACTAAATACATCATGGCATGATGTCCGC 5293
Db 5225 CCAGAATAGGTCACCTCACCCACCCCATAAACCAATACATCATGGCATGATGTCCGC 5284
Qy 5294 TGACCTGGAGGTGTCTCATAGCACTGGGTGTGGTAGGGGAGTCTTGTGACGCTTTGGC 5353
Db 5285 TGACCTGGAGGTGTCTCATAGCACTGGGTGTGGTAGGGGAGTCTTGTGACGCTCTGGC 5344
Qy 5354 CGCATACTGCTGACACAGGAGTGTGTCTATTGTGGCAGGATCATCTTGTTCGGGAA 5413
Db 5345 CGGTAATTGCTTGACAAACAGGAGTGTGTCTATTGTGGGTAGGATTTATTCTTGTTCGGGAG 5404
Qy 5414 GCCAGCTGTCTTCCCGACAGGGAAGTCTCTTACACGAGGTTTCGATGAGATGGAAGGTG 5473
Db 5405 GCGGGCCATTGTTCCCGACAGGGAGCTTCTTACACGAGGATTCGATGAAATGGAAGAGTG 5464

Qy 5474 TGCCTCAAACTTCTTACATCGAGCAGGGAATCGAGCTCGCCGAGCAATTTCAAGCAAAA 5533
Db 5465 GCGCTCGCACTTCTTACATCGAGCAGGGAATGAGCTCGCCGAGCAATTTCAAGCAGNA 5524
Qy 5534 GGCCTCGGTTGTGCAAAACGCGCAACAAAGCAGGAGCTGTCTCCCGTGGTGA 5593
Db 5525 AGCGCTCGGTTGTGCAAAACGCGCAACAAAGCAGGAGCTGTCTCCCGTGGTGA 5584
Qy 5594 GTCCAAGTGGGAGCCCTTGAGACTTCTGGGCGAAGCACAATGTGGAATTTTCAATCAGCGG 5653
Db 5585 GTCCAAGTGGGAGCCCTTGAGACTTCTGGGCGAAGCACAATGTGGAATTTTCAATCAGCGG 5644
Qy 5654 AATACAGTACCTAGCAGGCTTATCCACTCTCCCTGGAACCCCGGATAGCATCATTTGAT 5713
Db 5645 GATACAGTACTTAGCAGGCTTATCCACTCTCCCTGGAACCCCGGATAGCATCATTTGAT 5704
Qy 5714 GGCATTTACAGCTTTCTACTAGCAGGCTCACCACCAAAACACCTCTCTGTTTAAACAT 5773
Db 5705 GGCATTTACAGGCTTCTATCACCAGCCGCTCACCACCAAAAGTACCTCTCTGTTTAAACAT 5764
Qy 5774 CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGCGTCAGCTTTTCTGGG 5833
Db 5765 CTTGGGGGAGTGGGTGGCTGCCAACTCGCCCTCCCGCAGCGCTTCCGCTTTCTGGG 5824
Qy 5834 CGCGGCAATCGCGGAGCGGCTGTGGCAGCATAGGCTTGGGAAGTGTCTGTCGACAT 5893
Db 5825 CGCGGCAATCGCGGAGCGGCTGTGGCAGCATAGGCTTGGGAAGTGTCTGTCGACAT 5884
Qy 5894 CTTGGCGGCTTATGGGCGAGGCTAGCGGCGCATCTCGTGGCTTTTAAAGTTCATGAGCGG 5953
Db 5885 TCTGGCGGTTATGGAGCAGGAGTGGCGCGCTCGTGGCTTTTAAAGTTCATGAGCGG 5944
Qy 5954 CGAGTGCCTTCCACCGAGGACCTGGTCAACTTACTCCTGCCATCTCTCTCTCTGGTGC 6013
Db 5945 CGAGATGCCCTTCCACCGAGGACCTGGTCAATCTACTTCTGTCATCTCTCTCTCTGGGCG 6004
Qy 6014 CTTGTCGTCTGGGCTGCTGCGAGCAATACTGCGTGGCAGCTGGGCGCGGAGAGGG 6073
Db 6005 CTTGTCGTCTGGGCTGCTGCGAGCAATACTGCGTGGCAGCTGGGCGCGGAGAGGG 6064
Qy 6074 GGTGTGAGTGGATGAACCGGCTGATAGCTTCCCTTCGGGGGTAAACAGCTCTCCCC 6133
Db 6065 GGTGTGAGTGGATGAACCGGCTGATAGCTTCCCTTCGGGGGTAAATCATGTTTCCCC 6124
Qy 6134 TACGACATATGTGCTGAGAGGAGCTGACGACGCTGCTACTAGATCTCTCTAGCCT 6193
Db 6125 CACGCACTATGTGCTGAGAGCGACGCGCAGCGCTGTACTCAGATCTCTCTCAGCCT 6184
Qy 6194 TACCATCACTCAACTGCTGAAGCGGCTCCACAGTGGATTATAGAGGCTGCTCTACGCG 6253
Db 6185 TACCATCACTGAGCTGTGAAGGCTCCACAGTGGATTATAGAGGCTGCTCTCCACAC 6244
Qy 6254 ATGCTCCGCTCGTGGCTTAAAGGATGTTTGGGATTTGGATATGACGCTGTTGACTGACT 6313
Db 6245 GTGTTCCGCTCGTGGCTTAAAGGATGTTTGGGCTGGATATGACGCTGTTGACTGACT 6304
Qy 6314 AAGAAGTGGCTCAGTCCAAACTCCTCGCGGGTTACCGGAGTCCCTTTCTGTCTGTCATG 6373
Db 6305 CAAAGCTGGCTCCAGTCCAAAGCTCCTCGCGCAGCTACCTGGAGTCCCTTTTCTCTGTCG 6364
Qy 6374 CCAACGCGGTTACAGGAGGCTTGGCGGGGAGCGGCTATGTCGCAAAACACCTGCCCATG 6433
Db 6365 CCAACGCGGTTACAGGAGGCTTGGCGGGGAGCGGCTATGTCGCAAAACACCTGCCCATG 6424
Qy 6434 CGGAGCACAGATCGCGGACATGTCAAAAAACGGTTTCCATGAGGATCTGATAGGCTCTAGAAC 6493
Db 6425 TGGAGCACAGATCACCGGACATGTCAAAAAACGGTTTCCATGAGGATCTGTCGGGCTTAAGAC 6484
Qy 6494 CTGCGAGCAACAGTGGGCGAGGAACTTCCCATCAACCGCATACACACGGGACCTTGGAC 6553
Db 6485 CTGCGAGCAACAGTGGGCGAGGAACTTCCCATCAACCGCATACACACGGGCGGCTTGGAC 6544

QY 6554 ACCCTCCCGGCGCCCAACTATTTCAGGGGCTATGGCGGTGGCTGCTGAGGAGTACGT 5613
DB 6545 ACCCTCTCCAGGCCAAACTATTTCAGGGCGCTGTGGCGGTGGCGCTGAGGAGTACGT 5604
QY 6614 GGAGGTTACCGGTGGGGGATTTCCACTACGTGACGGGCATGACCACTGACAACTGAAA 6673
DB 6605 GGAGGTTACCGGGGTGGGGATTTCCACTACGTGACGGGCATGACCACTGACAACTGAAA 6664
QY 6674 GTGCCATGTCAGGTTCCGGCCCCCGAAATTTCTTACGGAGGTGGATGGAGTGGGTGCA 6733
DB 6665 GTGCCATGCCAGTTCCGGCTCTTGAAATTTCTTCTCGGAGGTGGACGGAGTGGGTGCA 6724
QY 6734 CAGGTACGCTCCGGCTGCAAACTCTTCTACGGGAGGAGCTGACGTTCCAGTCCGGCT 6793
DB 6725 CAGGTACGCTCCGGCTGCAAGGCTCTCTTACGGGAGGAGTTTACATTCAGGTCGGGT 6784
QY 6794 CAACCAATCTTGGTCCGGTCCGAGCTCCCATCGGAGCCCGAACCGGACGTAACTAGTCT 6853
DB 6785 CAACCAATACCTGGTTGGGTCA CAGCTTACCATCGGAGCCCGAACCGGATGTAGCAGTGTCT 6844
QY 6854 TACTTCCATGCTCACCGATCCCTCCCACTTACAGCAGAGACCGGCTAAGCTAGGCTGGC 6913
DB 6845 CACTTCCATGCTCACCGACCCCTCCCACTCA CAGCAGAAACGGCTAAGCGTAGGTTGGC 6904
QY 6914 TAGAGGTCTCCCGCTCTTTTACGAGCTCATCAGCTAGCCAGTTGTCTCGGCTTCTTT 6973
DB 6905 CAGGGGTCTCCCGCTCTTGGCCAGCTCTTCCAGCTAGCCAGTTGTCTCGGCTTCTTT 6964
QY 6974 GAAGGAGATGACTACCCACCATGACTCCCGGAGCGTGCCTCATCCGAGGCCAACCT 7033
DB 6965 GAAGGAGATGACTACCCACCATGCTCTCCGGAGCGCTGACCTCATCCGAGGCCAACCT 7024
QY 7034 CTTGTGCGCAGAGATGGCGGAAACATCACTCGCGTGGAGTCAGAGAAATAGGTAGT 7093
DB 7025 CTTGTGCGCAGAGATGGCGGGAACATCA CCGCGTGGAGTCGAGAAACAGGTGTGT 7084
QY 7094 AATTTGGACTCTTTGCAACCGCTTACGCGGAGGGGATGAGAGGAGATATCCGTCGC 7153
DB 7085 AGTCTGGACTCTTTGACCCGCTTCGAGCGGAGGAGTGAAGGGAAGTATCCGTTCC 7144
QY 7154 GCGGAGATCTGGGAAATTCAGGAAGTTCCCTCAGCGTTGCCATATGGGACGCCC 7213
DB 7145 GCGGAGATCTCGGGAATTCAGGAAGTTCCCGCAGCGATGCCCATCTGGGCGCGCCC 7204
QY 7214 GGACTAAATCTCCACTGTAGAGTCTTGAAGACCCCGACTACGTCCCTCCGCTGGT 7273
DB 7205 GGATTAACACCTCCACTGTTAGAGTCTTGAAGACCCCGACTACGTCCCTCCGCTGGT 7264
QY 7274 ACAGGATGCCCATTGGCACCTTCAAGGCTCTCTCAATACCACTCCACGGAGAAAGAG 7333
DB 7265 GCAGGGTGCCGTTGCCACTATCAAGGCCCTCTCAATACCACTCCACGGAGAAAGAG 7324
QY 7334 GACGGTTGCTGACAGAAATCAATGTCTTCTTCTGCTTGGCGAGCTCGGCCACTAAGAC 7393
DB 7325 GACGGTTGCTTAACAGAGTCTCCGCTGTCTTCTGCTTAGCGAGCTCGCTACTAAGAC 7384
QY 7394 CTTCCGTAGCTCCCGATCGTGGCGGTGTAGCGGCACCGCGACCGCCCTTCTGACCT 7453
DB 7385 CTTCCGCGAGCTCCGAAATCATCGGCGGTGACAGCGGCACCGCGACCGCCCTTCTGACCA 7444
QY 7454 GGCCTCCGACGACGCTGACAAAGGATCCGAGTTGAGTCTGACTCTCTCATGCCGCCCT 7513
DB 7445 GGCCTCCGACGACGCTGACAAAGATCCGAGTTGAGTCTGACTCTCTCATGCCGCCCT 7504
QY 7514 TGAAGGAGAGCCGGGGAACCCGATCTTCAGCGACGGGTCTTGGTCTACCGTGAAGGAG 7573
DB 7505 TGAGGGGAAACCGGGGACCCCGATCTCAGTGA CCGGTCTTGGTCTACCGTGAAGGAG 7564
QY 7574 GGCTAGTGAAGATGTCTGCTGCTGCTCAATGTCTTATACGTGGA CAGCGGCCCTGATCAC 7633
DB 7565 AGCTAGTGAAGATGTCTGCTGCTCAATGTCTTACATGGA CAGCGGCCCTGATCAC 7624
QY 7634 GCCATGCGCTCGGAGGAAAGTAAGCTGCCCATCAACCCGTTGAGCAACTCTTTGCTGCG 7693

DB 7625 GCATGCGCTCGGAGGAAAGAGCTGCCATCAACGGTTGAGCAACTCTTTGCTGCG 7684
QY 7694 TCACCACAACATGGTCTACGCCACAACATCCCGCAGCGCAAGCTCCGCGAGAAGGTT 7753
DB 7685 CCACCATACATGGTTTATGCCACAACATCTCGCAGCGAGGCTCGCGCAGAAGGTT 7744
QY 7754 CACCTTTGACAGTTCAAGTCTCGATGATCATTTACCGGGAGCTACTCAAGGAGATGAA 7813
DB 7745 CACCTTTGACAGCTGCAAGTCTCTGACGACCACTACCGGGAGCTGCTCAAGGAGATGAA 7804
QY 7814 GCGAAGGCGTCCACAGTTAAGCTTAACTTCTATCTATAGAGGAGCGCTGCAAGCTGAC 7873
DB 7805 GCGAAGGCGTCCACAGTTAAGCTTAACTTCTATCTCGTAGAAGAGCTGCAAGCTGAC 7864
QY 7874 GCGCCACAATTCGGCCAAATTCAAATTTGGCTATGGGGCAAGGAGCTCCGGAACTTATC 7933
DB 7865 GCGCCACAATTCGGCCAAATTCAAATTTGGCTATGGGGCAAGGAGCTCCGGAACTTATC 7924
QY 7934 CAGCAGGCGCTTAAACCATCCGCTCCGTGGGAGGACTTCTCGTAGACACTTGAAC 7993
DB 7925 CAGCAGGCGCTTAAACCATCCACTCCGTGGGAGGACTTCTCGTAGACACTTGAAC 7984
QY 7994 ACCAATTGACACCACCATCATGCAAAAAGTGAAGTTTCTCGCTCAACCCAGAGAAGG 8053
DB 7985 ACCAATTGACACCACCATCATGCAAAAAGTGAAGTTTCTCGCTCAACCCAGAGAAGG 8044
QY 8054 AGCCCAAGCAGCTCGCTTATCGTATTTCCAGACCTGGGAGTTGCTGTATGCGAGAA 8113
DB 8045 AGCCCAAGCAGCGCCCTTATCGTATTTCCAGATCTGGGAGTCCGTGTATGCGAGAA 8104
QY 8114 GATGGCGCTTTAGCAGGTGTCTCAACCTTCTCAGCGCGTGTATGGGCTCTCATACGG 8173
DB 8105 GATGGCGCTCTATGATGTGTCTCAACCTTCTCAGGCTGTATGGGCTCTCATACGG 8164
QY 8174 ATTTCAATCTCCCAAGCAGCGGTTCGAGTTCTCTGTTGAAATACCTTGGAAATCAAGAA 8233
DB 8165 ATTCCAGTACTCTCTGGGCGAGTCGAGTTCTCTGTTGAAATACCTTGGAAATCAAGAA 8224
QY 8234 ATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTTGACTCAACCGTCACTGAGAGTGA 8293
DB 8225 AAACCCCATGGGCTTTTATATGACACTCGCTGTTTCGACTCAACGGTCAACCGAGAACA 8284
QY 8294 CATTCGTTTGAAGGTCAAATTTCAATGTTGTGACTTGGCCCCCGAGGCCACAGGC 8353
DB 8285 CATTCGTTTGAAGGTCAAATTTTACCAATGTTGTGACTTGGCCCCCGAGGCCACAGGC 8344
QY 8354 CATAAGTCTGCTCACAGAGCGGCTTTTACATCGGGGTCCCTGACTAACTCAAAAGGCA 8413
DB 8345 CATAAAATCGCTCACAGAGCGGCTTTTATATCGGGGTCTCTGACTAACTCAAAAGGCA 8404
QY 8414 GAATCGGTTTATCGCGGTGCGCGCAAGTGGCGTCTGACGACTAGCTGCGGTAAATAC 8473
DB 8405 GAATCGGTTTATCGCGGTGCGCGGAGGCGTCTGACGACTAGCTGCGGTAAACAC 8464
QY 8474 CCTCACAATGTTACTTGAAGGCCACTGACGCTGTGAGCTGCAAAAGTCTCAGGACTGAC 8533
DB 8465 CCTCACAATGTTACTTGAAGGCCCTCTGACGCTGTGAGCTGCGAAGTCTCCAGGACTGAC 8524
QY 8534 GATGCTCGTGAACGGAGACGACTTGTGTTATCTGTGAAAGCGGGGAGCCAGGAGGA 8593
DB 8525 GATGCTCGTGAACGGAGACGACTTGTGTTATCTGTGAAAGCGGGGAGCCAGGAGGA 8584
QY 8594 TGCGGCGGCTTACGAGCTTACGAGGCTTATGACTAGGTATTTCCGCCCCCCCCCGGGGA 8653
DB 8585 CGCGCGGAGCTACGAGCTTTCAGGAGGCTATGACTAGGTACTTCGCGCCCCCCCCCGGGGA 8644
QY 8654 TCCGCCCCCAACCAATACGACTTGGAGCTGATAACATCATGTTCTCTCAATGTGTAGT 8713
DB 8645 CCGCGCCCCCAACCAATACGACTTGGAGCTGATAACATCATGTTCTCTCAATGTGTAGT 8704
QY 8714 CGCGCAGATGCTCTGGCAAAAGGATATCTACTCTACCCGCTGACCCACCCACCCCT 8773

```
Db 8705 CGCCACGATGATCAGGCAAAAGGGTGTACTACCTACCGGTGATCCACACACCCCT 8764
Qy 8774 TGCACGGCTGGTGGGAGACAGCTAGACACACTCCAATCAACTCTTGGCTAGGCAAT 8833
Db 8765 AGCAGGGCTGGTGGGAGACAGCTAGACACACTCCAGTTAACTCTGGCTAGGCAAT 8824
Qy 8834 CATCATGTATGCGCCACCCCTATGGCAAGGATGATCTGTAGTACTCACTTTTCTCCAT 8893
Db 8825 TATTATGTATGCGCCACTTTTGGGCAAGGATGATCTGTAGTACTCACTTTCTCCAT 8884
Qy 8894 CTTTCTAGCTCAAGGACAACTTGAAGAGCCCTGATGTGATGATCTAGGGCTTGTCTA 8953
Db 8895 CTTTCTAGCGCAGGAGCAACTTGAAGAGCCCTGACTGCCAGATCTACGGGGCTGTGA 8944
Qy 8954 CTCCTATTGAGCCTTGTGACTTACCTACCTCAGATCAATGAACGACTCCATGGTCTTAGGCGATT 9013
Db 8945 CTCCTATTGAGCCTTGTGACTTACCTCAGATCAATGAACGACTCCATGGCTTAGGCGATT 9004
Qy 9014 TACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAACT 9073
Db 9005 TTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAACT 9064
Qy 9074 TGGGTATCACCTTGGCAACTGAGACATCGGGCCAGAGTTCGGGCTTAAGTACT 9133
Db 9065 TGGGTATCACCTTGGCAACTGAGACATCGGGCCAGAGTTCGGGCTTAGGCTACT 9124
Qy 9134 GTCCAGGGGGGAGGGCGGCACCTTGTGGCAGATACCTTTAACTGGGCACTAAGAC 9193
Db 9125 GTCCAGGGAGGAGGGCGGCACCTTGTGGCAAACTACCTTTCAACTGGGCACTAAGAC 9184
Qy 9194 CAAGCTTAACTCACTCCAATCCGGCGCGCTCCAGCTGGACTTGTCTGGCTGGTTT 9253
Db 9185 CAACTTAACTCACTCCAATCCGGCGCGCTCCAGCTGGACTTGTCTGGCTGGTTT 9244
Qy 9254 CGCTGTTACAGCGGGGAGACATATATACAGCTGTCTGTGCGCGACCCGCTGGTT 9313
Db 9245 TGCTGGTTACAGCGGGGAGACATATATACAGCTGTCTGTGCGCGACCCGCTGGTT 9304
Qy 9314 TCCGTTGTGCTACTTCTTCTTGTAGGGTGGCATTTTACCTGCTCCCAACCGATG 9373
Db 9305 CATGCTGTGCTACTTCTTCTTGTAGGGTGGCATTTTACCTGCTCCCAACCGATG 9364
Qy 9374 AACGGGAGCTAACCACTCCAGCGCTTAAAGCCATTTCTTGTGTTTTTTTTTTT 9425
Db 9365 AACGGGAGATAACACTCCAGGCCAATAGGCCATCCCTCTTTTTTTTTTTT 9416

RESULT 5
US-08-904-686A-1
; Sequence 1, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSER: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 333..9362
; US-08-904-686A-1
```

```
Query Match 85.2%; Score 8170.4; DB 2; Length 9416;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 8636; Conservative 0; Mismatches 776; Indels 0; Gaps 0;

Qy 14 TGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAACTACTGTCTTCCAGCAAA 73
Db 5 TGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTGTCTTCCAGCAAA 64
Qy 74 AGCGTCTAGCCATGGCGTTAGTATGAGTGTGTCAGCTCCAGGACCCCTCCCGG 133
Db 65 AGCGTCTAGCCATGGCGTTAGTATGAGTGTGTCAGCTCCAGGACCCCTCCCGG 124
Qy 134 AGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATTGCCAGACACCGGTC 193
Db 125 AGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATTGCCAGACACCGGTC 184
Qy 194 TTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGCGTGTCCCGCGAGACTGTCTAG 253
Db 185 TTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGCGTGTCCCGCGAGACTGTCTAG 244
Qy 254 CCGAGTAGTGTGGTCCGAAAGCCCTTGTGTTACTGCTGATAGGGTGTCTGCGAGTG 313
Db 245 CCGAGTAGTGTGGTCCGAAAGCCCTTGTGTTACTGCTGATAGGGTGTCTGCGAGTG 304
Qy 314 CCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACGAATCTTAACCTCAAGAAAAAC 373
Db 305 CCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACGAATCTTAACCTCAAGAAAAAC 364
Qy 374 CAAACGTAAACCAACCCCGCCCAACAGGAGCTCAAGTTCCCGGGCGGTGTGTCAGTCT 433
```

Db	365	CAAACGTAAACAAACCGCGGCCCAAGGACGTCAAGTTCCCGGGCGGTGGTCAAGATCGT	424
Qy	434	TGTTGAGATTTACCTGTTCGCGCAGGGGCCCCAGGTTGGGTGTCGCGCGACTAGGAA	493
Db	425	TGTTGAGATTTACCTGTTCGCGCAGGGGCCCCAGGTTGGGTGTCGCGCGCCAGGAA	484
Qy	494	GGTTTCGAGCGGTGCGAACTCTGTGGAAGGCGCAACCTATCCCAAGGCTCGCGACC	553
Db	485	GACTTCGAGCGGTGCGAACTCTGTGGAAGGCGCAACCTATCCCAAGGCTCGCGGCC	544
Qy	554	CGAGGCGAGGCTCGGCTCAGCCCCGGGTACCTTGGCCCTCTATGGCAATCAGGCGCT	613
Db	545	CGAGGCGAGGCTCGGCTCAGCCCCGGGTACCTTGGCCCTCTATGGCAATCAGGCGCT	604
Qy	614	GGGTTGGGCAAGTGGCTCCTGTACCCCCGGGCTCCCGGCTAGTTGGGGCCCAACGGA	673
Db	605	AGGTTGGGCAAGTGGCTCCTGTACCCCCGGGCTCCCGGCTAGTTGGGGCCCAACGGA	664
Qy	674	CCCCGGGCTAGTTCGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGC	733
Db	665	CCCCGGGCTAGTTCGGTAAAGTTCATCGATACCCCTTACATGCGGCTTCGC	724
Qy	734	CGATCTCAGGGGTACATTCGGTTCGTCGGGCGCCCTAGGGGGCGCTGCCAGGGCCCT	793
Db	725	CGATCTCAGGGGTACATTCGGTTCGTCGGGCGCCCTAGGGGGCGCTGCCAGGGCCCT	784
Qy	794	GGCACCGGTGTCGGGTTCTGGAGGACGGCGTGAACATATGCAACAGGGAACCTTGC	853
Db	785	GGCACATGGTTCGGGTTCTGGAGGACGGCGTGAACATATGCAACAGGGAACCTTGC	844
Qy	854	TTGCTCTTTCTATCTTCCTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	913
Db	845	TTGCTCTTTCTATCTTCCTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	904
Qy	914	TTATGAGTGGCAACGCTGCGGGATATACCATGTCACGAACGACTGCTCCAACTCAAG	973
Db	905	TTACGAAGTGCACAACTGTCGGGATATATCATGTACGAACGACTGCTCCAACTCAAG	964
Qy	974	CATTGTGTAGGACGCGGACGTGATCATGCTACTCCCGGGTCCGTCGCTGCTGCTGCTA	1033
Db	965	CATTGTGTAGGACGCGGACTTGATCATGCTACTCCCGGGTCCGTCGCTGCTGCTGCTG	1024
Qy	1034	GGAGGTTAACAGTTCCTGCTGGGTAGCGCTCACTCCCAACGCTCGCGGCGAGGAATGC	1093
Db	1025	GGAGGCAACTCTCTCCGCTGCTGGGTAGCGCTCACTCCCAACGCTCGCGGCGAGGA	1084
Qy	1094	CAGGTCCTCCACTAGCAATAACGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1153
Db	1085	CACCATCCCAACCAACGACGATACGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCT	1144
Qy	1154	CTGCTCCGCTATGATGCTGGGGATCTCTGGGATCTATTTTCTCTCTCCAGCTGTT	1213
Db	1145	CTGTTCCGCTATGATGCTGGGGATCTCTGGGATCTGTTTCTCTCTCTCTCTCTCTCT	1204
Qy	1214	CACCTTCTCGCTCGCGGCAATGACAGTGCAGGACTGCAACTGCTCAATCTATCCCGG	1273
Db	1205	CACCTTCTCGCTCGCGGCAATGACATACAGGACTGTAACCTGCTCAATTTATCCCGG	1264
Qy	1274	CCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGATGATGATGATGATGAT	1333
Db	1265	CCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGATGATGATGATGATGAT	1324
Qy	1334	CCTAGTGGGTGCGAGTTCCTCCGATCCCAACAGCTGCTGTCGACATGCTGGCGGGGC	1393
Db	1325	CCTAGTGGGTGCGAGTTCCTCCGATCCCAACAGCTGCTGTCGACATGCTGGCGGGGC	1384
Qy	1394	CCACTGGGGAGTCTCGCGGCTTTCCTACTATTTCCATGGTGGGAACTGGGCTAAGGT	1453
Db	1385	CCACTGGGGAGTCTCGCGGCTTTCCTACTATTTCCATGGCGGGAACTGGGCTAAGGT	1444
Qy	1454	TCTGATGTGCGCTACTCTTTTGGCGGCTTGAACGGGGAGACCCACACGACGGGAGGGT	1513
Db	1445	TCTGATGTGCTACTCTTTTGGCGGCTTGAACGGGGATACCCACGCTGACAGGGGGGC	1504

Qy	1514	GGCGGCGCACACACCTCCGGGTTACGCTCCCTTTCTCATCTGGGCGCTCTCAGAAAAT	1573
Db	1505	GCAAGCCAAAACCAACACAGGCTCGTGTCCATGTTTCGAAGTGGGCCGCTCTCAGAAAAT	1564
Qy	1574	CCAGCTTTGTAATACCAACGCGAGCTGGCACATCAACAGAGACTGCGCTTAAATTTGCAATGA	1633
Db	1565	CCAGCTTATTAACCAATGGGNGTTGGCACATCAACAGAGACTGCGCTTAAATTTGCAATGA	1624
Qy	1634	CTCCCTCCAAAATCGGGTTCTTTGCGCGCTGTTTAAACGACACAAGTTTAACTCTGTCGG	1693
Db	1625	CTCTCTCCAGACTGGGTTCTTTGCGCGCTGTTTAAACACACATAGTTTCAACTCTGTCGG	1684
Qy	1694	GTGCGCGGAGCGATGCGGCGCTGCGGCGCTGTTGACTGTTGCGCGCGAGGTTGGGCGCC	1753
Db	1685	GTGCGCGAGCGATGCGGCGCTGCGGCGCTGTTGACTGTTGCGCGCGAGGTTGGGCGCC	1744
Qy	1754	CATCACTTACTTAAGCCTAAACAGCTCGGATCAGAGGCGCTTATTTGCTGGCATTAACGCGCC	1813
Db	1745	CATTACTTATCTAGCTCTAGCAGATCAGACAGAGGCGCATATTTGCTGGCACTACCCACC	1804
Qy	1814	TCGACCGTGTGGTGTCTGTTACCCCGCTCGCAGGTGTGTGTCAGTGTATTTGTTTCACCCC	1873
Db	1805	TCCACAATGTACCATCGTACCTCGTTCGGAGGTGTGCGGCCAGTGTACTGCTTCACCCC	1864
Qy	1874	AAGCCCTGTTGTTGGTGGGACCAACGATCGTTCGGGTGCTCCCTACGTATAGCTGGGGGA	1933
Db	1865	AAGCCCTGCTGCTGGGACGACCGATCGTTCGGGTGCTCCCTACGTATAGTGGGGGA	1924
Qy	1934	GAATGAGACAGAGTGTGCTCTCAACAAACACGCGTCCGCCCAACAGGCAACTGGTTCGG	1993
Db	1925	GAACGAGACTGACGCTGCTGCTCTCAACAAACACGCGGCCCGCGAGGCAACTGGTTCGG	1984
Qy	1994	CTGTACATGGATGAATAGTACTCGGTTCTACTAAGAGCTGCGGAGGTCCCGTGTAAACAT	2053
Db	1985	CTGCACATGATGAATAGCACCGGTTTACAAAGACATGTGGGGGGCCCCCGTGTAAACAT	2044
Qy	2054	CGGGGGGTCGGTAAACCGGACCTTGATCTGCGGCCACCGGACTGCTTCGGGAGCACCCCGA	2113
Db	2045	CGGGGGGTCGGCAACAAACCCCTGACCTGCGGCCACCGGACTGCTTCGGGAGCACCCCGA	2104
Qy	2114	GGCTACTTACACAAAATGTGGCTCGGGGCTCGGTTGACACCTAGGTGCTCTAGTAGACTA	2173
Db	2105	GGCTACTTACACAAAATGTGGTTCGGGGCTTGGCTGACACCTAGGTGCTCTAGTAGACTA	2164
Qy	2174	CCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCTTTTAAAGTTAGGAT	2233
Db	2165	TCCATACAGGCTCTGGCATTTACCCCTGCACTGTTTAACTTTTACCATCTTCAAGTTAGGAT	2224
Qy	2234	GTATGTGGGGGCTGGAGCACAGGCTCAATGCGCGCATGCAATTTGGACTCGAGGAGCG	2293
Db	2225	GTATGTGGGGGCTGGAGCACAGGCTCAATGCTGTCATGCAATTTGGAGCCCGAGGAGCG	2284
Qy	2294	CTGTAACTTCGAGACAGGATAGGTGCAAGCTCAGCCCGCTGCTGCTCTCAACACAGA	2353
Db	2285	TTGTACTTCGAGACAGGATAGGCGCGAGCTCAGCCCGCTGCTGCTCTCAACACAGA	2344
Qy	2354	GTGSCAGATCTGCCCTGCTGCTTACACACCTACCGGCTTTATCCACTGGTTTGAATCCA	2413
Db	2345	GTGSCAGATCTGCCCTGCTTCTTCTCACACCTTACAGCTCTGCTCCACTGGCTTGAATCCA	2404
Qy	2414	TCTCATCAGAACATCGTGGACGTGCAATACCTGTACGGGTAGGGTACGGGTTTGTCTC	2473
Db	2405	CCTCCATCAGAACATCTGGGACGTGCAATACCTATACGGTATAGGGTCAGCGGTTGTCTC	2464
Qy	2474	CTTTGCAATCAATGGGAGTACATCTGTTGCTTTTCTCTCTCTGCGACGCGCGGT	2533
Db	2465	CTTTGCAATCAATGGGAGTATGTTCTGTTGCTTTTCTCTCTCTGCGACGCGCGGT	2524
Qy	2534	GTGTGCTGCTGTTGGGATGATGCTGTATAGCCAGGCTGAGGCGCGCTTAGAGAACTT	2593
Db	2525	CTGTGCTGCTGTTGGATGATGCTGTATAGCCAGGCGCGCGCTTTGGAGAACTT	2584

QY 2594 GGTGTCCTCAATGGCGGTCGGTGGCCGGAGCGCATGTTATCTCTCTCTTTCTGTGT 2653
Db |||||
2595 GGTGTCCTCAATTCGGCGTCTGTGGCCGGCGCATGCGCATCTCTCTCTTTCTGTGT 2644
QY 2654 CTTCTGCGCGCTGGTACATTAAGGGCAGGTGGCTCTGGGGGGGGGTATGCTTTT 2713
Db |||||
2645 CTTCTGTGCGCGCTGGTACATCAAAAGGAGGCTGGTCCCTGGGGCGCATATGCTCTTT 2704
QY 2714 TGGCGTATGGCGCTGCTCTGCTCTCTACTGGCGTTACCAACGAGCTTACGGCTTGG 2773
Db |||||
2705 TGGCGTATGGCGCTGCTCTGCTCTTGTGGCATTAACCGCGAGCTTACGCCATGGA 2764
QY 2774 CCGGAGATGGCTGATCGTGGGGGTGCGGTTCTTGTAGGTCTGTGTTATCTTGACCT 2833
Db |||||
2765 CCGGAGATGGCTGATCGTGGGAGGCGGTTTGTGGTCTGGTACTCTGACTTT 2824
QY 2834 GTCACCATACTACAAGTGTCTCACTAGGCTCATATGGTGGTGTACAAATCTTTATCAC 2893
Db |||||
2825 GTCACCATACTACAAGTGTCTCGCTAGGCTCATATGGTGGTGTACAAATTTTACCA 2884
QY 2894 CAGAGCGAGCGGCATGCAAGTGTGGTCCCCCGCTCAACGTTGGGAGGCGCGGA 2953
Db |||||
2885 CAGAGCGAGCGGACTTACATGTGTGATCCCCCGCTCAACGCTCGGGAGGCGCGGA 2944
QY 2954 TGCCATCATCTCTCAGCTGTGCGGTTCAATCAGAGTTAAATTTTGACATCACCAACT 3013
Db |||||
2945 TGCCATCATCTCTCATGTGGCAGTCCATCCAGAGTAATCTTTGACATCACCAACT 3004
QY 3014 CTTGCTCGGCATPACTCGGCGGCTCATGTGCTCCAGGCTGGCATAAAGAGTGGCCGTA 3073
Db |||||
3005 TCTAATGTCATPACTCGGTCGCTCATGTGCTCCAAGCTGGCATAAAGAGTGGCCGTA 3064
QY 3074 CTTGCTGGCGCTCAAGGCTCATTCGTGCATGATGTAGTCGAAAGTCGCGGGG 3133
Db |||||
3065 CTTGCTGGCGCTCAAGGCTCATTCATGTCATGTAGTCGAAAGTGGTGGGG 3124
QY 3134 TCATTATGTCCAAATGGTCTTCATGAAGCTGGGCGGCTGACAGGTAGTACGTTTATAA 3193
Db |||||
3125 TCATTATGTCCAAATGGCTTCATGAAGTGGGCGGCTGACAGGCAGTACATTTACAA 3184
QY 3194 CCATCTTACCCCACTGCGGACTGGGCCCCACGCGGCTTACGAGACTTTGGCGTGGCGGT 3253
Db |||||
3185 CCATCTTACCCCGCTACGGGATTTGGCCACGCGCGGCTTACGAGACTTTGCGGTGGCAGT 3244
QY 3254 AGAGCCGCTGCTTCTCGCCCATGGAGACCAAGGTATCATCTGCGGAGCAGACACCGC 3313
Db |||||
3245 GGAGCCGCTGCTTCTCGCATAGGACCAAGATCATCTGCGGAGCAGACACCGC 3304
QY 3314 TGCGTGTGGGACATCATCTTGGGCTTACCCGCTCCGCGGAAAGGAGATATT 3373
Db |||||
3305 GGCGTGTGGGACATCATCTTGGGCTTGGCGCTTCCGCGGAAAGGAGATATT 3364
QY 3374 TTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTTGGGACTCTTTGCGCCCATCACGCG 3433
Db |||||
3365 CTTGCGCCCGGCGGATAGTCTTGAAGGGCGGGGTTGCGACTCTCTCGCGCCCATCACGCG 3424
QY 3434 CTACTCCCAACAAACGCGGGGCTACTTGGTTCATCATCTAGCCTCAGAGCGCGGA 3493
Db |||||
3425 CTACTCCCAACAAACGCGGGGCTACTTGGTTCATCATCTAGCCTTACAGGCGCGGA 3484
QY 3494 CAAGAACAGGTTCGAAGGGGAGGTTCAAGTGGTTTCTACCGCAACACATCTTCTCTGCG 3553
Db |||||
3485 CAAGAACAGGTTCGAAGGGAGGTTCAAGTGGTTTCCACCGCAACACATCTCTCTGCG 3544
QY 3554 GACCTGCATCAACGCGCTGTGGACTGTCTACCATGGCGCTGGCTCGAAGACCTTAGC 3613
Db |||||
3545 GACCTGCATCAACGCGCTGTGGAACGTTTACCATGGTGTGGCTCAAGACCTTAGC 3604
QY 3614 CCGTCCAAAGGTTCNAATCACCAATGTACACCAATGTAGACCTGGACCTGTGGCTG 3673
Db |||||
3605 CCGCCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACGAGGCTCTGTGGCTG 3664
QY 3674 GCAGGCGCCCCCGGGCGGCTCCATGACACCATGACCTGTGGCAGCTCGGACCTTTA 3733

Db 3665 GCCCAAGCCCCCGGGCGGTTCTTTGACACCATGACCTGTGGCAGCTCAGACCTTTA 3724
QY 3734 CTTGGTCAACAGACATGCTGATGTCATTTCGGTGCCTGGCGGAGCCACAGCAGGGGAAG 3793
Db 3725 CTTGGTCAACAGACATGCTGACGTCATTCCGGTGCCTGGCGGAGCCACAGTAGGGGAG 3784
QY 3794 TCTACTCTCCCCCAGGCGGCTCTCTACTCTGAAAGGCTCTCTCGGGTGGTCCATTGCTTTG 3853
Db 3785 CTTGCTCTCCCCCAGGCTGTCTCTACTTTGAAGGCTCTCTCGGGTGGTCCACTGCTCTG 3844
QY 3854 CTTTTCGGGCAAGTCGTGTGGCGTCTTTTCGGGCTGTGTGTGCAACCGGGGGTTCGGAA 3913
Db 3845 CTTTTCGGGCAAGTCGTGTGGGCTCTTTCCGGCTGCGGTATGCAACCGGGGGTTCGGAA 3904
QY 3914 GCGGTTGGACTTCATACCGGTTGAGTCTATGGAACACTACCATGCGGTCTCCGGTCTTCAC 3973
Db 3905 GCGGTTGGACTTTGTGTCGGGTAGAGTCCATGGAACACTACTATGCGGTCTCCGGTCTTCAC 3964
QY 3974 AGACAACTCAACCCCCCGGCTGTACCGCAGACATTCGAAGTGGCAGATCTGCAACGCTCC 4033
Db 3965 GGACAACTCATCCCCCGGCGGTACCGAGTCATTTCAAGTGGCCCACTTACAGCTCC 4024
QY 4034 TACTGGCAGCGGCAAGAGCACAAAGTGCCTGGTGTGCTATGCAACCCAAAGGTTCAAGGT 4093
Db 4025 CACTGGCAGCGGCAAGAGTACTAAAGTGCCTGGCTGCATATGCAGCCCAAGGTTCAAGGT 4084
QY 4094 GCTCGTCTGNAACCGGTCGTTGCGGCCACCTTAGGGTTTGGGGGTATATGTCNAAGGC 4153
Db 4085 GCTCGTCTCAATCCGTCGTTGCGGTACCTTAGGGTTTGGGGGTATATGTCNAAGGC 4144
QY 4154 ACAGGTATCGACCTTAACATCAGAACTGGGTTAAGGACCAATTACACAGCGCGGCTCCAT 4213
Db 4145 ACAGGTATTCACCCCAACATCAGAACTGGGTTAAGGACCAATTACACAGCGCGGCTCC 4204
QY 4214 TACGTACTCCACTPATGGCAAGTTCCTTGGCGACGCTGGTGTGTTCTGGGGGCGCTATGA 4273
Db 4205 CACATCTCTACTATGGCAAGTTTCTTGGCGATGGTGGTGTCTGGGGGCGCTTATGA 4264
QY 4274 CATCATATATGATGATGCGCACTCAACTGACTCGACTACCATCTTGGGCACTCGGCAC 4333
Db 4265 CATCATATATGATGATGCGCACTTCAACTGACTCGCACTCAATCTTGGGCACTCGGCAC 4324
QY 4334 AGTCTTGACCAAGCGAGACGGCTGGAGCGCGCTCGTCTGTCTGCGCACCGCTACACC 4393
Db 4325 AGTCTTGACCAAGCGAGACGGCTGGAGCGCGCTGTGCTGTCTGCGCACCGCTACGCC 4384
QY 4394 TCCGGGATCGGTTTACCGTGGCCACACCCCAATPATCGAGGAAATAGGCTGTCCAAACATGG 4453
Db 4385 TCCGGGATCGGTCACCGTGGCCACACCCCAACATCGAGGAGGTGGCCCTGTCTAATCTTGG 4444
QY 4454 AGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGCGCATCAAGGGGGGAGGCACTCT 4513
Db 4445 AGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGCGCATCAAGGGGGGAGGCACTCT 4504
QY 4514 CATTTCTGCCATTTCCAAGAGAAATGTGACAGCTCGCCGCAAGCTGACAGGCGCTCGG 4573
Db 4505 CATTTCTGTCTTCCAAGAAAGTGGCAGCAGCTCGCCGCAAGCTGTGAGGCTCGG 4564
QY 4574 ACTGAACGCTGTAGCATATTTACCGGGGCTTGTGTGTCGTCTATPCGCCCTATTCGGAGA 4633
Db 4565 AATCAACGCTGTGGGATTTACCGGGGCTCGATGTGTCTCATACCACTATTCGGAGA 4624
QY 4634 CGTGTGTGTGGCAACAGACGCTCTAATGACGGGTTTACCGGCGATTTTGACTCAGT 4693
Db 4625 CGTGTGTGTGGCAACAGACGCTCTGTATGACGGGCTTATCGGGCGACTTTTGACTCAGT 4684
QY 4694 GATCGACTGCAATCATGTGTCAACCGACAGTGCAGTTTCACTTGGATCCCACTTCCAC 4753
Db 4685 GATCGACTGTAACATGTGTCAACCGACAGTGCAGTTTGAATCCCACTTCCACCTTCCAC 4744
QY 4754 CATTGAGACGACGACCGGTGCCCCAAGACGCGGTGTGCGCTCGCAACCGCGGAGGTAGAAC 4813

Db 4745 CATTGAGACGACGACCGTCTCAAGACGACGAGTGTGCGGCTCGCAGCGCGGGGTAGGAC 4804
QY 4814 TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAACGGCCCTCGGGCAT 4873
Db 4805 TGGCAGGGGTAGGAGAGGCATCTACAGGTTTGTGACTCCGGGAGAACGGCCCTCGGGCAT 4864
QY 4874 GTTCGATTTCTTCGGTCTGTGTGAGTGTCTATGACGGGGCTGTGCTTGTGTATGAGTCTAC 4933
Db 4865 GTTCGATTTCTTCGGTCTGTGTGAGTGTCTATGACGGGGCTGTGCTTGTGTATGAGTCTAC 4924
QY 4934 GCCGCTGAGACCTTCGGTGTAGTGTGCGGCTTACCTAAATATACACACGAGGTTGCGGCTGTG 4993
Db 4925 CCGGGCGGAGACCTTCGGTGTAGTGTGCGGCTTACCTGAAACACACAGGTTGCGGCTGTG 4984
QY 4994 CCAGGACCACTCTGGAGTTCTGGGAGAGCGTCTTTCACAGGCTCACCCACATAGATCCCA 5053
Db 4985 CCAGGACCACTTCGGAGTTCTGGGAGAGTGTCTTTCACAGGCTCACCCATATAGATGCAC 5044
QY 5054 CTTCTGTCTCCAGACTAAACAGGAGGAGACAACTTTCTTACTCTGGTGGCATATCAAGC 5113
Db 5045 CTTCTGTCTCCAGACTAAACAGGAGGAGACAACTTTCTTACTCTGGTGGCATATCAAGC 5104
QY 5114 TACAGTGTGCGGCGAGGCTCAAGCTCCACCTCCATCGTGGGACCAAAATGTGGAAGTGTCT 5173
Db 5105 CACGGTGTGCGGCGAGGCTCAGGCCCCACCTCCATCATGAGGATCAAAATGTGGAAGTGTCT 5164
QY 5174 CATAAGCTGAACCTACACTGACAGGCGGCAACACCCCTCTGTATAGGCTAGGAGCGGT 5233
Db 5165 CATAAGCTGAACCTACAGCTGACAGGCGGCAACACCCCTCTGTATAGGCTAGGAGCGGT 5224
QY 5234 CCAAAATGAGGTCACTCTACACACCCCATAACTAAATATCATATGCGCATGTCTGGC 5293
Db 5225 CCAGATGAGTCACTCTACACCCCATAACTAAATATCATATGCGCATGTCTGGC 5284
QY 5294 TGACCTGAGGTGCTCACTAGCACCTGGTGTAGGCGGAGTCTCTGACGTTTGGC 5353
Db 5285 TGACCTGAGGTGCTCACTAGCACCTGGTGTAGGCGGAGTCTCTGACGTTTGGC 5344
QY 5354 CGCATCTGCTGACGACGAGCGTGTGCTATTGTGGGAGGATCATCTTTGTCGGGAA 5413
Db 5345 CGCGTATTGCTGACACAGGAGTGTGCTATTGTGGGTAGGATTAATCTTTGTCGGGAG 5404
QY 5414 GCCAGCTGTCTGCTCCGACAGGGAAGTCTCTTACAGGAGTTCGATGAGATGGAAGAGTG 5473
Db 5405 GCCGGCCATTGTTCCCGACAGGAGCTCTCTACAGGAGTTCGATGAATGGAAGAGTG 5464
QY 5474 TGCCTCACACTTCTTACATCGAGAGGGAATCGAGCTGCCAGCAATTCAGCAAAA 5533
Db 5465 CGCCTCGCACCTCCCTTTACATCGAGCAGGGAATCGAGCTGCCAGCAATTCAGCAAAA 5524
QY 5534 GCGCTCGGGTTGTGCAACCGGCCACCAAGACGAGGCTGCTGCTCCCGTGGTGA 5593
Db 5525 AGCGCTCGGGTTACTGCAACAGGCCACCAAGACGAGGCTGCTGCTCCCGTGGTGA 5584
QY 5594 GTCCAAGTGGCGAGCCCTTGAGACCTTCTGGGCGAAGCACATGTGGAATTTTCATCAGCGG 5653
Db 5585 GTCCAAGTGGCGAGCCCTTGAGACCTTCTGGGCGAAGCACATGTGGAATTTTCATCAGCGG 5644
QY 5654 AATACAGTACTAGCAGGCTTATCACTCTGCTGGAAACCCCGGATAGCATATTGAT 5713
Db 5645 GATACAGTACTAGCAGGCTTATCACTCTGCTGGAAACCCCGGATAGCATATTGAT 5704
QY 5714 GGCATTTACAGTCTTATCACTAGCCGCTCACACCCCAAAACACCCCTCTGTTTAAAT 5773
Db 5705 GGCATTTACAGTCTTATCACTAGCCGCTCACACCCCAAAACACCCCTCTGTTTAAAT 5764
QY 5774 CTTGGGGGATGGGGTGTCCCAACTCGCTCTCCAGCGCTGGTTCAGCTTTTCGTTGGG 5833
Db 5765 CTTGGGGGATGGGGTGTCCCAACTCGCTCTCCAGCGCTGGTTCAGCTTTTCGTTGGG 5824
QY 5834 CGCCGGCATCGCCGAGGGGCTGTGTGGCAGCATAGGCCCTTGGGAAGGTGCTGAGGACAT 5893
Db 5825 CGCCGGCATCGCCGATGGGCTGTGTGGCAGCATAGGCCCTTGGGAAGGTGCTGAGGACAT 5884

QY 5894 CTTTGGCGGCTATGCGGCGAGGGGTAGCGGGCAGCTCGTGGCCTTTAAGTTCATGAGCGG 5953
Db 5885 TCTGCGGGTTATGAGCAGGAGTGGCGGCGCTCTGTTGGCTTTTAAAGTTCATGAGCGG 5944
QY 5954 CGAGTGTCCCTCCACCGAGGACCTGCTCAACTTACTTCCCTGCCATCTCTCTCTCTGGTGC 6013
Db 5945 CGAGTGTCCCTCCACCGAGGACCTGCTCAACTTACTTCTCTGCCATCTCTCTCTCTGGCGC 6004
QY 6014 CTTGTGCTGCGGGTCTGTGTGCGAGCAATACTCGTGGCAGCTGGGCGCGGAGAGG 6073
Db 6005 CTTGTGCTGCGGGTCTGTGTGCGAGCAATACTCGTGGCAGCTGGGCGCGGAGAGG 6064
QY 6074 GGTGTGAGTGTGATGACCGGCTGATAGCTTTCGCTTCGCGGGTAAACACGCTCTCCCC 6133
Db 6065 GGTGTGAGTGTGATGACCGGCTGATAGCTTTCGCTTCGCGGGTAAATCATGTTTCCCC 6124
QY 6134 TACGCACTATGTCCCTGAGAGCGCTGCAGCAGTGTCTCTCAGATCCTCTCTAGCCT 6193
Db 6125 CAGCACTATGTCCCTGAGAGCGACCGCGAGCGGTGTACTCAGATCCTCTCTCAGCCT 6184
QY 6194 TACCACTACTCAACTGTCTGAAGCGCTCCACAGTGGATTAATGAGGAGTGTCTTAGGCC 6253
Db 6185 TACCACTACTCAACTGTCTGAAGCGCTCCACAGTGGATTAATGAGGAGTGTCTTAGGCC 6244
QY 6254 ATGCTCGGCTCGTGGTAAAGGATGTTTGGGATTCGATATGACAGCTGTGTGACTGACTT 6313
Db 6245 GTTCTCGGCTCGTGGTAAAGGATGTTTGGGACTGGATATGACCGGTGTGTGACTGACTT 6304
QY 6314 CAAAGCTCGCTCAGTCCAACTCCTGCGCGGTTTACCGGAGTCCCTTTCTGTCTATG 6373
Db 6305 CAAAGCTCGCTCAGTCCAACTCCTGCGGAGTACCTGGAGTCCCTTTTCTCGTG 6364
QY 6374 CCAACCGGGTACAAGGAGTCTGCGGGGGAGCGGATCATGCAAAACACCTGCCCATG 6433
Db 6365 CCAACCGGGTACAAGGAGTCTGCGGGGGAGCGGATCATGCAAAACACCTGCCCATG 6424
QY 6434 CGAGCAGATCGCGGACATGTCAAAACGTTCCATGAGGATCGTAGGGCTAGAAC 6493
Db 6425 TGAGCAGCAGATCACCGGACATGTCAAAACGTTCCATGAGGATCGTAGGGCTAGAAC 6484
QY 6494 CTCAGCAACAGCTGGCAGGAACTTCCCATCAACGCAATACACACGAGACCTTGAC 6553
Db 6485 CTGAGCAACAGCTGGGATGGAATTTCCCATCAACGCAATACACACGAGGCGCTGAC 6544
QY 6554 ACCCTCCCGGCGCCAACTATTTCAGGGCGCTATGCGGGTGGTCTGAGGAGTACGT 6613
Db 6545 ACCCTCTCAGCGCCAACTATTTCAGGGCGCTATGCGGGTGGTCTGAGGAGTACGT 6604
QY 6614 GAGGTTACGCGTGGGGATTTCCACTAGTGAAGGCGATGACCACTGACACGCTAAA 6673
Db 6605 GAGGTTACGCGTGGGGATTTCCACTAGTGAAGGCGATGACCACTGACACGCTAAA 6664
QY 6674 GTGCCATGCAAGTTCGCGCGCCGAAATTTTTCAGGAGTGGATGAGTGGCGTTGCA 6733
Db 6665 GTGCCATGCAAGTTCGCGCGCTCTCTGAAATTTCTCTGAGGTTGAGCAGGAGTGGCA 6724
QY 6734 CAGTACGCTCCGCGGTGCAAACTCTTCTACGGGAGGAGTCTACGTTTCCAGGTCGGGCT 6793
Db 6725 CAGTACGCTCCGCGGTGAGGCTCTCTCTACGGGAGGAGTCTACATTTCCAGGTCGGGCT 6784
QY 6794 CAACCAATATCTTGGTGGAGTCCCATGATGCGAGCGGCAACCGGACGTAACAGTGTCT 6853
Db 6785 CAACCAATATCTTGGTGGAGTCCCATGATGCGAGCGGCAACCGGACGTAACAGTGTCT 6844
QY 6854 TACTTCCATGCTACCGATCCCTCCCATATPACAGCAGAGCGGCTAAGCGTAGGCTGGC 6913
Db 6845 CACTTCCATGCTACCGACCCCTCCCATACAGCAGAAACCGGCTAAGCGTAGGCTGGC 6904
QY 6914 TAGAGGCTCTCCCGCTCTTTTAGCAGCTCATCAGCTAGCCAGTGTGTCTGCGGCTCTTTT 6973
Db 6905 CAGGGGCTCTCCCGCTCTTTTAGCAGCTCTTTAGCAGCTCTTTAGCAGCTCTTTAGCAGCTCTTTT 6964

QY 6974 GAAGCGACATGCACTACCCACCATGACTCCCGAGCGCTGACCTCATCGAGCGCAACCT 7033
DB 6965 GAAGCGACATGCACTACCCACCATGCTCTCCGAGCGCTGACCTCATCGAGCGCAACCT 7024
QY 7034 CTTGTGGCGGAGAGATGGGCGGAAACATCACTCGGTGAGTGCAGAAATAGAGTAGT 7093
DB 7025 CTTGTGGCGGAGAGATGGGCGGAAACATCACCGCGTGGAGTCGGAGAACAAAGTGGT 7084
QY 7094 AATTCTGACCTCTTTCGAAACCGCTTCACGCGAGGGGATGAGAGGAGATATCCGTGCG 7153
DB 7085 AGTCTGAGCTCTTTCGAAACCGCTTCGAGCGAGGAGATGAGAGGAGATATCCGTGTC 7144
QY 7154 GCGGAGATCTTCGAAACATCCAGAAAGTTCCCTTCAGCGTTGCCCATATGGGACGCGCC 7213
DB 7145 GCGGAGATCTTCGAAACATCCAGAAAGTTCCCCGACGGATGCCATCTGGGCGGCGCC 7204
QY 7214 GGACTACAATCTCTCACTGCTAGAGTCTCTGAAGAACCCGAGCTACAGTCCCTCCGGTGGT 7273
DB 7205 GGATTACAACCTCTCACTGTTAGAGTCTCTGAAGAACCCGAGCTACGTCCTCCGGTGGT 7264
QY 7274 ACAGGATGCCATTTGCCACCTACCAAGGCTCTCCAATACCACTCCACGAGAAAGAG 7333
DB 7265 GCAGGGTGCCGTTGCCACCTATCAAGGCCCTCCAATACCACTCCACGAGAAAGAG 7324
QY 7334 GACGGTTGCTGACAGAAATCCAAATGTGTCTCTGCTTGGGGAGCTCGCACCTAAGAC 7393
DB 7325 GACGGTTGCTTAAACAGAGTCTCTCGTGTCTTCTGCTTTCGCGAGCTCGCTACTAAGAC 7384
QY 7394 CTTCCGTAGCTCCGATCGTCGCGGTTGATAGCGGCAACCGGACCGCCCTTCTCTGACCT 7453
DB 7385 CTTCCGAGCTCCGAATCATCGGCCGTGACAGCGGCAACCGGACCGCCCTTCTCTGACCA 7444
QY 7454 GGCCTCCGACGAGCTGACAAAGATCCGAGTTGAGTCTGACTCTCCATGCCCCCTCT 7513
DB 7445 GGCCTCCGACGAGCTGACAAAGATCCGAGCTTGAAGTCTGACTCTCCATGCCCCCTCT 7504
QY 7514 TGAAGGGAGCGGGGACCCGATCTCAGCGACGGGTCTTGGTCTACCGTGAAGTGAAGA 7573
DB 7505 TGAGGGGAAACCGGGGACCCGATCTCAGTACCGGGTCTTGGTCTACCGTGAAGTGAAGA 7564
QY 7574 GGCTAGTGAGGATGTCGTGCTGCTCAATGCTTATAGTGGACGAGCGCCCTGATCAC 7633
DB 7565 AGCTAGTGAGGATGTCGTGCTGCTCAATGCTTACATGACAGCGGCGCTTGTATCAC 7624
QY 7634 GGCATCGCTCGGAGGAAAGTAAAGTGCCCATCAACCCGTTGAGCAACTCTTCTGCTGG 7693
DB 7625 GGCATCGCTCGGAGGAAAGTAAAGTGCCCATCAACCGCTTGAAGCAACTCTTCTGCTGG 7684
QY 7694 TCACCAACATGTTCTACGCCACAACATCCCGCAGCGCAAGCCCTCCGCAAGAAAGT 7753
DB 7685 CCACCATACATGGTTTATGCCACAACATCTCGCAGCGCAGGCGCTTGGCAGAAAGT 7744
QY 7754 CACCTTTGACAGATGCAAGTCTCGATGATCATTTACCGGAGCTACTCAAGGAGATGAA 7813
DB 7745 CACCTTTGACAGCTGCAAGTCTCGACGACCACTACCGGAGCTGCTCAAGGAGATGAA 7804
QY 7814 GCGAAGGGCTCCACAGTTAAGGCTTAAGCTTCTATATAGAGGAGGCTGCAAGCTGAC 7873
DB 7805 GCGAAGGGCTCCACAGTTAAGGCTTAACTCTATATCCGTTAGAGGAGGCTGCAAGCTGAC 7864
QY 7874 GCGCCCATTCGCGCAATCCAAATTTGGCTATGGGCAAGGAGCTCCGCAACCTATC 7933
DB 7865 GCGCCCATTCGCGCAATCCAAATTTGGCTATGGGCAAGGAGCTCCGCAACCTATC 7924
QY 7934 CAGCAGGGCGGTTAAACCAATCCGCTCCGTTGGGAGGACTTGTCTGGAAGACACTGAAAC 7993
DB 7925 CAGCAGGGCGGTTAAACCAATCCGCTCCGTTGGGAGGACTTGTCTGGAAGACACTGTAAC 7984
QY 7994 ACCAATTGACACCATCATGGCAAAAGTGAAGTTTCTCGCTCCAAACGAGAGAGGG 8053
DB 7985 ACCAATTGACACCATCATGGCAAAAGTGAAGTTTCTCGCTCCAAACGAGAGAGGG 8044
QY 8054 AGGCCGCAAGCGAGCTCGCCTTATCGTATTTCCAGACCTGGGAGTTTCGTGATCGAGAA 8113

DB 8045 AGGCGTAAGCCAGCCCGCTTATCGTATTTCCAGATCTGGAGTCCGTGTATCGAGAA 8104
QY 8114 GATGGCCCTTTACGACGTGCTCCACCCTTCTCAGCGGTGATGGCTCTCATACGG 8173
DB 8105 GATGGCCCTTTATGATGCTCTCCACCCTTCTCAGGTGATGGCTCTCATACGG 8164
QY 8174 ATTTCAATACTCCCCCAAGCAGCGGTTCGAGTTCCTGCTGAATACCTTGGAAATCAAGAA 8233
DB 8165 ATTCAGTACTCTCTCGGCGAGCTGAGTTCCTGCTGAATACCTTGGAAATCAAGAA 8224
QY 8234 ATGCCCTATGGCTTCTCATATGACACCCGCTGTTTGTGACTCAACCGTCACTGTAGAGTGA 8293
DB 8225 AAACCCCATGGCTTTTCATATGACACTCGCTGTTTCGACTCAACCGTCAACGAGAACGA 8284
QY 8294 CATTCGTTGTGAGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGAGCAGACAGC 8353
DB 8285 CATCCGTTGTGAGAGTCAATTTACCAATGTTGTGACTTGGCCCCCGAGCAGACAGC 8344
QY 8354 CATAAGTTCGCTCACAGAGCGCTTTATATCGGGGTCCTCTGACTAATCAAAAGGCA 8413
DB 8345 CATAAATTCGCTCACAGAGCGCTTTATATCGGGGTCCTCTGACTAATCAAAAGGCA 8404
QY 8414 GAACCTGGGCTTATCGCCGTCGCCGCAAGTGGGTGCTGACGACTAGCTCGGTAAATAC 8473
DB 8405 GAACCTGGGCTTATCGCCGTCGCCGCGAGCGGCTGCTGACGACTAGCTCGGTAAACAC 8464
QY 8474 CCTCACATGTTACTTGAAGCCACTGACGCTCTGAGCTCCAAAGCTCCAGGACTGAC 8533
DB 8465 CCTCACATGTTACTTGAAGCCCTCTGAGCCTCTGAGCTCGAAGCTCCAGGACTGAC 8524
QY 8534 GATCTCGTGAACGAGACGACCTTGTGCTTATCTGTGAAGCGGGAAACCCAGAGAGA 8593
DB 8525 GATCTCGTGAACGAGACGACCTTGTGCTTATCTGTGAAGCGGGAAACCCAGAGAGA 8584
QY 8594 TGCAGCGGCTTACGAGCCTTTCAGGAGCTTATGACTAGTATTTCCGCCCCCCCCGGGA 8653
DB 8585 CGCGCGAGCCTACGAGTCTTTCAGGAGCTTATGACTAGTACTCCGCCCCCCCCGGGA 8644
QY 8654 TCCGCCCCAACAGAAATACGACCTGGAGCTGATTAACATCATGTTCTCCAATGTTGCTAGT 8713
DB 8645 CCGCCCCAACAGAAATACGACTTGGAGCTGATTAACATCATGTTCTCCAATGTTGCTAGT 8704
QY 8714 CGCGCAGTCACTCGCAAAAGGTTATCTACTCAACCGTGACCCACCCACCCCTCT 8773
DB 8705 CGCCACGATGCATCAGCAAAAGGTTGTTACTCTACCCGTTGATCCACCCACCCCTCT 8764
QY 8774 TGCACGGCTCGTGGGAGACAGCTAGACACACTCCAATCAACTCTTGGCTAGGCAATAT 8833
DB 8765 AGCAGGGCTCGTGGGAGACAGCTAGACACACTCCAAGTTAACTCTTGGCTAGGCAACAT 8824
QY 8834 CATCATGATCGGCCACCTTATGGGCAAGGATGATTTCTGATGACTCACTTTTCTCCAT 8893
DB 8825 TATTATGATGCGGCCACTTTTGTGGCAAGGATGATTTCTGATGACTCACTTTCTTCTCCAT 8884
QY 8894 CTTTCTAGCTCAAGAGCAACTTGAAGGCGCTGATTTGTGATGACTCACTACGGGCTTGTGA 8953
DB 8885 CTTTCTAGCGGAGCAACTTGAAGGCGCTGATTTGCCAGCTGACAGATCTAGGGGCTGTTA 8944
QY 8954 CTCCATTGAGCCTTGAACCTTACCTCAGATCATTTGAACGACTCCAATGGTCTTTAGGCGCAT 9013
DB 8945 CTCCATTGAGCCTTGAACCTTACCTCAGATCATTTGAACGACTCCAATGGTCTTTAGGCGCAT 9004
QY 9014 TACACTCCAGTACTCTCTCAGGTTGAGATCAATAGGTTGGCTTTCATGCTCAGGAAACT 9073
DB 9005 TTCACTCCATGTTACTCTCCAGGTTGAGATCAATAGGTTGGCTTTCATGCTCAGGAAACT 9064
QY 9074 TGGGTTACACCTTTCGAACTCGAGACATCGGCGAGAGTGTCCGCTTAAGTACT 9133
DB 9065 TGGGTTACACCTTTCGAGTCTGAGACATCGGCGCAGGAGCGTCCGCTTAGGCTACT 9124
QY 9134 GTCCAGGGGGGAGCGGCCACTTGTGGCAGATACCTCTTTTAACTGGGCAAGTAAAGGAC 9193

Db 9125 GTCCAGGGAGGGCGCCACCTGTGGCAATACCTCTTCACTGGCAGTAAAC 9184
Qy 9194 CAAGCTTAAACTCACTCCCAATCCCGCGCGCTCCAGCTGGACTTGTCTGGCTGGTTCTG 9253
Db 9185 CAAACTTAAACTCACTCCCAATCCCGCTGCGTCCCGCTGGACTTGTCTGGCTGGTTCTG 9244
Qy 9254 CGCTGTTACAGGGGGGAGACATATATACAGCCTGTCTCGTCCCGGACCCCGCTGGTT 9313
Db 9245 TGTGTTGTACAGGGGGGAGACATATATACAGCCTGTCTCGTCCCGGACCCCGCTGGTT 9304
Qy 9314 TCGTTGTGCTACTCTCTACTTCTGTAGGGGTAGGCATTACTGTCTCCCAACCGATG 9373
Db 9305 CATGCTGTGCTACTCTCTACTTCTGTAGGGGTAGGCATTACTGTCTCCCAACCGATG 9364
Qy 9374 AACGGGGAGCTAACCACTCCAGGCCTTAAGCCATTTCTCTGTTTTTTTTTTT 9425
Db 9365 AACGGGGAGATAAACACTCCAGGCCAATAGGCCATCCCCCTTTTTTTTTTTT 9416

RESULT 6
US-09-315-850-1
; Sequence 1, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FURE, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeLland, Le-Nhung

REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 900703G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: CDS
LOCATION: 333...9362
US-09-315-850-1

Query Match 85.2%; Score 8170.4; DB 3; Length 9416;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 8636; Conservative 0; Mismatches 776; Indels 0; Gaps 0;

Qy 14 TGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATCTGTCTTCAAGCAGAA 73
Db 5 TGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATCTGTCTTCAAGCAGAA 64
Qy 74 AGCGTCTAGCCATGGCGTTAGTATGATGTCTGTGACGCTCCAGGACCCCTCCCGG 133
Db 65 AGCGTCTAGCCATGGCGTTAGTATGATGTCTGTGACGCTCCAGGACCCCTCCCGG 124
Qy 134 AGAGCCATAGTGTCTGCGAACCGGTGAGTACACCGGAATTCGCCAGGACCGGGTCC 193
Db 125 AGAGCCATAGTGTCTGCGAACCGGTGAGTACACCGGAATTCGCCAGGACCGGGTCC 184
Qy 194 TTTCTTTGGATCAACCCGCTCAATGCTGGAGATTGGGGGTGCCCGCGAGACTGCTAG 253
Db 185 TTTCTTTGGATCAACCCGCTCAATGCTGGAGATTGGGGGTGCCCGCGAGACTGCTAG 244
Qy 254 CGAGTAGTGTGGGTGCGAAAGGCTTGTGTGCTGCTGCTGATAGGTTGCTTGGAGTG 313
Db 245 CGAGTAGTGTGGGTGCGAAAGGCTTGTGTGCTGCTGCTGATAGGTTGCTTGGAGTG 304
Qy 314 CCCGGGAGGTCTGTAGACCGTGACCATGACGACGAATCTTAAACCTCAAGAAAAAAC 373
Db 305 CCCGGGAGGTCTGTAGACCGTGACCATGACGACGAATCTTAAACCTCAAGAAAAAAC 364
Qy 374 CAAACGTAAACCAACCCGCGCCACAGGACGTCAAGTTCCCGGGGGTGGTCAAGTCGT 433
Db 365 CAAACGTAAACCAACCCGCGCCACAGGACGTCAAGTTCCCGGGGGTGGTCAAGTCGT 424
Qy 434 TGTGAGATTACCTGTTCCGCGCAGGGGCCCCAGGTTGGGTGTGCGCGGACTAGAA 493
Db 425 TGTGAGATTACCTGTTCCGCGCAGGGGCCCCAGGTTGGGTGTGCGCGGCCCCAGGAA 484
Qy 494 GGCTTCGAGCGGTGCGAACCTCGTGAAGCGACACCTATCCCAAGGCTCCCGGACC 553
Db 485 GACTTCGAGCGGTGCGAACCTCGTGAAGCGACACCTATCCCAAGGCTCCCGGACC 544
Qy 554 CGAGGCGAGGCGCTTGGGCTCAGCCCGGGTACCCCTTGGCCCTCTATGCGCAATAGGCGCT 613
Db 545 CGAGGCGAGGCGCTTGGGCTCAGCCCGGGTACCCCTTGGCCCTCTATGCGCAATAGGCGCT 604
Qy 614 GGGGTGGGCGAGGATGGCTCTGTCAACCCCGGGCTCCCGGCTAGTTGGGGCCCCACGGA 673
Db 605 AGGTGGGCGAGGATGGCTCTGTCAACCCCGGGCTCCCGGCTAGTTGGGGCCCCACGGA 664
Qy 674 CCCCCGCGTAGGTGCGGTAACCTGGGTAAGGTCACTCGATACCTTACATCGCGCTTCGC 733
Db 665 CCCCCGCGTAGGTGCGGTAACCTGGGTAAGGTCACTCGATACCTTACATCGCGCTTCGC 724
Qy 734 CGATCTCATGGGGTACATTCGGCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCTT 793
Db 725 CGATCTCATGGGGTACATTCGGCTCGTGGCGCCCCCTTAGGGGGCGCTGCCAGGCGCTT 784

Qy	794	GGCACAGGTGTCGGGTTCTGGAGGACGGCTGAACATATGCAACAGGGAACTTGCCCGG	853
Db	785	GGCACATGTTGTCGGGTTCTGGAGGACGGCTGAACATATGCAACAGGGAATCTGCCCGG	844
Qy	854	TTGGCTCTTTCTCTATCTTCCTCTTGGCTCTGCTGCTCTGTTTGCACCATCCACGCTTCCGC	913
Db	845	TTGCTCTTTTTCTATCTTCTCTTTGGCTCTGCTGCTCTGCTGACCAACCACGCTTCCGC	904
Qy	914	TTATGAAGTGGCAACGTGTCCGGATATACCATGTCACGAAGCATGTCTCCAACCTCAAG	973
Db	905	TTACGAAGTGCACAACGTGTCCGGATATATCATGTCACGAACGATGTCTCCAACGCAAG	964
Qy	974	CATTGTGTATGAGCAGCGAGCATGATCATGCAATCTCCCGGTGCGTCCCTGTGTTCA	1033
Db	965	CATTGTGTATGAGCAGCGAGCTTGATCATGCACTACTCTCGGTGCGTCCCTGCGTCTG	1024
Qy	1034	GGAGGTTAAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACAGCTCGCGCACAGGAATGC	1093
Db	1025	GGAAAGCAACTCTCTCCCGTCTGGGTAGCGCTCACTCCACAGCTCGACGCCAGGAACGT	1084
Qy	1094	CAGCGTCCCACTACGACAATACAGACGCCACGTGCACTTGCTCGTTGGGACGGCTGCTTT	1153
Db	1085	CACCATCCCAACACGACGATACAGCCACGTCGATCTGCTGTTGGGCGGCTGCTTT	1144
Qy	1154	CTGCTCCGCTATGTACGTGGGGGATCTCTGGGATCTATTTTCTCTGCTCTCCAGCTGTT	1213
Db	1145	CTGTTCCGCTATGTACGTGGGGACCTCTGCGGATCTGTTTCTCTGCTCTCAGCTGTT	1204
Qy	1214	CACCTTCTCGCCTCGCGGCATGACAGGTGCAGGACTGCAATGCTCTCAATCTATCCCGG	1273
Db	1205	CACCTTCTCGCCTCGCGGCATGTGCAATTACAGGACTGTAACTGCTCAATTTATCCCGG	1264
Qy	1274	CCATGTATCAGGTCAACCGCATGGCTTCGGGATATGATGATGAACTGGTTCACCTACAAACAGC	1333
Db	1265	CCATGTGTCCGTTACCGTATGGCTTCGGACATGATGATGAACTGGTTCGCCCAACACAGC	1324
Qy	1334	CCTAGTGGTTCGAGTTGCTCCGGATCCCAACAGCTGTCTGTGGACATGTGCGCGGGGC	1393
Db	1325	CCTAGTGGTTCGAGTTACTTCCGGATCCCAACAGCCGTCTGTGACATGTTGCGCGGGGC	1384
Qy	1394	CCACTGGGGAGTCTTGGGGGCTTGCCTACTATTTCATGGTAGGGAACCTGGGCTAAGGT	1453
Db	1385	CCACTGGGGAGTCTTGGCGGCTTGGCTACTATTTCATGGCGGGGAACCTGGGCTAAGGT	1444
Qy	1454	TCTGATTGTGGCGCTACTCTTTTGGCGGCTTGACGGGGAGACCCACACGACGGGAGGGT	1513
Db	1445	TCTGATTGTGATGTACTTTTGTGGGTTGACGGGGATACCCACGTGACGGGGGGGC	1504
Qy	1514	GGCCGGCCACACCACTCCGGTTTACGTCCTCTTTTCTCATCTGGGCGCTCTCAGAAAAT	1573
Db	1505	GCAAGCCAAAACCAACAGGCTCGTGTCCATGTTGCAAGTGGGCGCTCTCAGAAAAT	1564
Qy	1574	CCAGTTGTGAATACCAACGCACTGGCACATCAACAGGACTGCCCTTAATTTGCATATGA	1633
Db	1565	CCAGCTATAAACCAACCAATGGGAGTTGGACATCAACAGGACTGCCCTGAACTGCAATGA	1624
Qy	1634	CTCCCTCCAAACTGGGTTCTTTGCGCGCTGTTTTAGCGCACACAAGTTCAACTCGTCCGG	1693
Db	1625	CTCTCTCCAGACTGGGTTCTTTCGCGCTGTTCTACACATAGTTTTCACTGCTCCGG	1684
Qy	1694	GTGCCCGAGCGCATGGCAGCTCCCGCCCAATTGACTGGTTCCGCCAGGGGTGGGGCCC	1753
Db	1685	GTGCCAGAGCGCATGGCCAGTGCAGCACCATTTGACAAGTTCCGACCAGGGATGGGGTCC	1744
Qy	1754	CATCACCTATATTAAGCCTAAACGCTCGGATCAGAGGCTTATTTGCTTGGCATTTACGGGCC	1813
Db	1745	CATTACTTATCTGAGCTTAGCATCAGACAGAGGCCATATTTGCTGGCATACTCCACC	1804
Qy	1814	TCGACCGTGTGGTGTGCTPACCCGGTTCGACAGTGTGTGGTCCAGTGTATTTTTCACCCC	1873
Db	1805	TCCCAATGTACCATCTGCTGTGGAGGTGTGGGCCAGTGTACTGCTTTCACCCC	1864
Qy	1874	AAGCCCTGTTGTGTGGGGACCAACCGATCTGTTCCGGTGTCTCCTACGTATATAGCTGGGGGA	1933

[illegible]

Db 2945 TGCCATATCCCTCATGTGCGCAGTCCATCCAGAGCTAATCTTTTGACATCACCAAACT 3004
QY 3014 CTGTGCGCATACCTGCGCCGCTCATGTGCTCCAGGCTGSCATAAAGAGAGTCCCGTA 3073
Db 3005 TCTAATGCGCATACCTGCGCCGCTCATGTGCTCCAGGCTGSCATAAAGAGTCCCGTA 3064
QY 3074 CTTTGTGCGCGCTCAAGGCTCATGTGCTCATGTGCTGCGGAAAGTCCCGGGG 3133
Db 3065 CTTTGTGCGCGCTCAAGGCTCATGTGCTCATGTGCTGCGGAAAGTCCCGGGG 3124
QY 3134 TCATTATGTCAAATGCTTCTCATGAAGCTGGGCGGCTGACAGGTACGTATTATAA 3193
Db 3125 TCATTATGTCAAATGCTTCTCATGAAGCTGGGCGGCTGACAGGCACGTACATTATACAA 3184
QY 3194 CCATCTTACCCCACTCGCGGACTGGGCCACCGGSCCTACGAGACTTTCCGCTGCGGT 3253
Db 3185 CCATCTTACCCCGCTACGGGATTTGGCCACGCGCGGCTTACGAGACTTTGCGGTGCGAT 3244
QY 3254 AGAGCCGCTGCTTCTCCGCCATGAGACCAAGGTTCATCACCCTGGGGAGCAGACCCG 3313
Db 3245 GGAGCCGCTGCTTCTCCGACATGAGACCAAGATTCATCCTGGGGAGCAGACCCG 3304
QY 3314 TGCGTGTGGGACATCATCTTGGTCTTACCCGCTCCGCGGAGGAGGAGATATT 3373
Db 3305 GGCGTGTGGGACATCATCTTGGTCTTCCCGCTCCGCGGAGGAGGAGATATT 3364
QY 3374 TTTGGGACCGGTGATGCTTCAAGGCAAGGCTGGGCTCTTCCGCGCATCACGGC 3433
Db 3365 CTTGGGCGCGCGCTAGTCTTGAAGGCGGGGTTGCGACTCTCCGCGCATCACGGC 3424
QY 3434 CTACTCCCAACAAAGCGGGGCTTGTGCTCATCATCTAGCTTACAGGCGGGG 3493
Db 3425 CTACTCCCAACAGCGGGGCTTGTGCTCATCATCTAGCTTACAGGCGGGG 3484
QY 3494 CAAGAACAGGTGAAAGGAGGTTCAAGTGGTCTTACCGCAACACAATCTTTCTGCG 3553
Db 3485 CAAGAACAGGTGAGGAGGTTCAAGTGGTCTTCAACGCAACACAATCTTTCTGCG 3544
QY 3554 GACCTGCATCAAGCGGTGCTGGAATGCTTACCTAGCTGGGCTGCTGGAAGCCCTAGC 3613
Db 3545 GACCTGCTCAAGCGGTGCTGGAATGCTTACCTAGCTGGGCTTCAAGAGCCTTAGC 3604
QY 3614 CGGTGCAAAAGGTCCATACCCAAATGTACCAATGTAGACTGACCTCGTGGCTG 3673
Db 3605 CGCGCAAAAGGCGCAATACCCAGATGTACCTAATGTGGACAGGACTCGTGGCTG 3664
QY 3674 GCAGGCGCGCGCGGCGCTCCATGACACCATGAGCTGTGGCAGCTCGGACCTTTA 3733
Db 3665 GCCAAGCGCGCGCGGCGCTTCTTGACACCATGACCTGTGGCAGCTCAGACCTTTA 3724
QY 3734 CTTGCTCAGGACATGTGATGTCAATTCGCGTGGCGGCGAGCGGAGCGGGAAG 3793
Db 3725 CTTGCTCAGGACATGTGATGTCAATTCGCGTGGCGGCGGCGGAGCAGTGGGGAG 3784
QY 3794 TCTACTCTCCCGGCGGCTCTCTACTGAAAGGCTCTCGGCTGCTCAATGCTCTTG 3853
Db 3785 CTTGCTCTCCCGGCGGCTCTCTACTGAAAGGCTCTTCCGCTGCTGCTCACTGCTCTG 3844
QY 3854 CTTTGGGCGACGCTGCGGCTCTTCCGCGCTGCTGTCGACCGCGGGGCTCGGAA 3913
Db 3845 CTTTGGGCGACGCTGCGGCTCTTCCGCGCTGCTGTCGACCGCGGGGCTTGGGAA 3904
QY 3914 GCGGTGACCTTATACCGTGTAGTCTATGGAATCTACCATGCGGCTCTCGGCTCTTAC 3973
Db 3905 GCGGTGACCTTGTGCGCGGTAGAGTCCATGGAATCTACTATGCGCTCTCGGCTCTTAC 3964
QY 3974 AGACAACTCAACCCCGCGGCTGTACCGCAGACATTCCAAAGTGGCACTATGACGCTCC 4033
Db 3965 GGACAACTCATCCCCCGCGGCTGTACCGCAGTCAATTTCAAGTGGGCCACCTACAGCTCC 4024
QY 4034 TACTGGCAGCGGCAAGAGCACAAAGTGGCGGCTGCTGATGAGCCCAAGGTTACAGGT 4093
Db 4025 CACTGGCAGCGGCAAGAGTACTTAAAGTGGCGGCTGCATATGAGCCCAAGGTTACAGGT 4084

QY 4094 GCTGCTCTGAACCCGCTCCGTTGCGGCCACCTTAGGTTTGGGGGTATATGTCCAAGGC 4153
Db 4085 GCTGCTCTCAATCCGTTGCGCTTACCTTAGGTTTGGGGGTATATGTCTTAAGGC 4144
QY 4154 ACAGGTATCGACCCCTAATCATCAGAACTGGGTAAGGACCAATTACACGGGCGGCTCCAT 4213
Db 4145 ACAGGTATCGACCCCTAATCATCAGAACTGGGTAAGGACCAATTACACAGGCGCCCGT 4204
QY 4214 TACGTACTCCACTATGCAAGTTCTCTTGGCGAAGTGGTGGTCTTGGGGGCGCTATGA 4273
Db 4205 CACATACTCTACTATGCAAGTTCTCTTGGCGAAGTGGTGGTCTTGGGGGCGCTATGA 4264
QY 4274 CATCATATATGTGATGCTGCTCACTGACTCGACTCGACTCGACTCGACTCGACTCG 4333
Db 4265 CATCATATATGTGATGCTGCTCACTGACTCGACTCGACTCGACTCGACTCGACTCG 4324
QY 4334 AGTCTCGACCAAGCGGAGAGCGGCTGGAGCGGCTCGTGGTCTCGGCGCTACACCG 4393
Db 4325 AGTCTCGACCAAGCGGAGAGCGGCTGGAGCGGCTTGTGCTGCTCGGCGCTACGCG 4384
QY 4394 TCCGGGATCGGTTACCGTCCCAACCCCAATATCGAGGAAATAGGCTGTCCAAATGG 4453
Db 4385 TCCGGGATCGGTTACCGTCCCAACCCCAATATCGAGGAGGTGGCTGTCTAATACTGG 4444
QY 4454 AGAGATCCCTTCTATGCAAAAGCCATCCCAATTTGAGGCCATCAAGGGGGGAGGCTCT 4513
Db 4445 AGAGATCCCTTCTATGCAAAAGCCATCCCAATTTGAAAGCCATCAGGGGGGAGGCTCT 4504
QY 4514 CATTTCTGCCATTCGAAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG 4573
Db 4505 CATTTCTGTCAATTCGAAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCTCGG 4564
QY 4574 ACTGAAGCTGTAGCATATTTACCGGGGCTTGTATGTGCTCGCTCATACCGCTATCGGAG 4633
Db 4565 AATCAAGCTGTGGCTATTTACCGGGGCTTGTATGTGCTCGCTCATACCAACTATCGGAG 4624
QY 4634 CGTGTGTGTGTGGGCAACAGACGCTCTAATGACGGGTTTACCGGGGATTTTGAAGTCACT 4693
Db 4625 CGTGTGTGTGTGGGCAACAGACGCTCTGATGACGGGCTATACCGGGGACTTTGACTCAGT 4684
QY 4694 GATCGACTGCAATATGCTGTACCCAGACAGTTCGACTTCGACTTCGACTTCGACTTCAC 4753
Db 4685 GATCGACTGCAATATGCTGTACCCAGACAGTTCGACTTCGACTTCGACTTCGACTTCAC 4744
QY 4754 CATTGAGACGACACCGTCCCGCAAGACGCGGTGCGCTCGCAACGCGGAGTAGAAC 4813
Db 4745 CATTGAGACGACACCGTCCCGCAAGACGCGGTGCGCTCGCAACGCGGAGTAGAAC 4804
QY 4814 TGGCAGGGTAGGAGTGGCATCTACAGTTTGTGACTCCAGGAGAACCGGCTCGGGCAT 4873
Db 4805 TGGCAGGGTAGGAGTGGCATCTACAGTTTGTGACTCCCGGAGAACCGGCTCGGGCAT 4864
QY 4874 GTTCGATTCTTCGCTCTGCTGTGAGTGTATGACGCGGCTGTGCTTGGTATGAGCTCAC 4933
Db 4865 GTTCGATTCTTCGCTCTGCTGTGAGTGTATGACGCGGCTGTGCTTGGTATGAGCTCAC 4924
QY 4934 GCGCGCTGAGACCTCGGTTAGTGTGGGCTTACCTAAATACACAGGTTGCGGCTGTG 4993
Db 4925 CCGCGCGGAGACCTCGGTTAGTGTGGGCTTACCTAAATACACAGGTTGCGGCTGTG 4984
QY 4994 CCAGGACCATCTGGAGTTCTGGGAGGCTTCTTCAAGGCTTCAACCCACATAGATGCCA 5053
Db 4985 CCAGGACCATCTGGAGTTCTGGGAGGCTTCTTCAAGGCTTCAACCCACATAGATGCCA 5044
QY 5054 CTTCTGTCCAGACTAAACAGGCGAGGAGCAACTTTCTTCTTCTGTGGGCAATATCAAGC 5113
Db 5045 CTTCTGTCCAGACTAAACAGGCGAGGAGCAACTTTCTTCTTCTGTGGGCAATATCAAGC 5104
QY 5114 TACAGTGTGCGGCGGCTCAAGTCTCACTTCCATCTGTGGGCAACCAATGTGGAGTGTCT 5173
Db 5105 CACGCTGTGCGGCGGCTCAAGGCTTCCATCTCATGGGATCAAAATGTGGAGTGTCT 5164

QY 5174 CATAAGGCTGAAACCTACATCTGACGGGCCAACACCCCTGCTGTATAGGCTAGGAGCGGT 5233
DB 5165 CATAAGGCTGAAACCTACGCTGCAAGGCCAACACCCCTTGTGTACAGGCTGGAGCGGT 5224
QY 5234 CAAAATGAGGTATCTCTCACACACCCCATAACTAAATACATCATATGCGATCATGTGCGC 5293
DB 5225 CCAGAATGAGGTACACCTCACACACCCCATAAACCAATACATCATATGCGATCATGTGCGC 5284
QY 5294 TGACCTGAGGTGCTGCTACATAGCACCTGGTGTGTAGCGGAGTTCCTTGACGCTTTGGC 5353
DB 5285 TGACCTGAGGTGCTGCTACATAGCACCTGGTGTGTAGCGGAGTTCCTTGACGCTTCGCG 5344
QY 5354 CGCATCTGCTGACGACAGGAGGTGTGTCTATTGTGGCGAGGATCATCTTGTGCGGAA 5413
DB 5345 CGCGTATTGCCCTGACAAACAGGAGGTGTGTCTATTGTGGTAGGATTATCTTGTCCGGAG 5404
QY 5414 GCCAGCTGTGCTTCCGACAGGGAAGTCTCTACACAGAGTTCGATGAGATGGAAGAGT 5473
DB 5405 GCGGCGCATTTGTTCCGACAGGAGGCTTCTCTACAGGAGTTCGATGAAATGGAAGAGT 5464
QY 5474 TGCCCTACAACTTCCTTACATCGACGAGGAATGAGCTCGCGAGCAATTCAGACAAA 5533
DB 5465 GCGCTCGACCTCCCTTACATCGACGAGGAATGAGCTCGCGAGCAATTCAGACAAA 5524
QY 5534 GCGCTCGGCTTGTGCAACGCGCACCAAGCAAGCGAGGCTGTCTGCCGTGGTGA 5593
DB 5525 AGCGCTCGGCTTGTGCAACAGCCACCAACAGCGAGGCTGTCTGCCGTGGTGA 5584
QY 5594 GTCCAAGTGGAGGCTTTGAGACCTTCTGGCGAAGCACATGTGGAATTTATCATCAGCG 5653
DB 5585 GTCCAAGTGGAGGCTTTGAGACATTTCTGGCGAAGCACATGTGGAATTTATCATCAGCG 5644
QY 5654 AATACAGTACCTAGCAGGCTTATCACTCTGCTGGAACCCCGATAGCATATTGAT 5713
DB 5645 GATACAGTACTTAGCAGGCTTATCACTCTGCTGGAACCCCGATAGCATATTGAT 5704
QY 5714 GGCATTACAGCTTCTATCACTAGCGGCTCACACCCAAACACCTCTGTTTAAAT 5773
DB 5705 GGCATTACAGCTCTATCACAGCGGCTCACACCCAAAGTACCTCTGTTTAAAT 5764
QY 5774 CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCTCCAGCGCTGGCTCAGCTTTCTGGG 5833
DB 5765 CTTGGGGGATGGGTGGCTGCCAACTCGCCCCCGCGCTTCGGCTTTCTGGG 5824
QY 5834 CGCGGCTCGCGGAGCGGCTGTGGAGCATAGGCTTGGGAAGGTCTCGTGGACAT 5893
DB 5825 CGCGGCTCGCGGAGCGGCTGTGGAGCATAGGCTTGGGAAGGTCTGTGGACAT 5884
QY 5894 CTTGGCGGCTATGGGCGAGGAGTACCGGCGCACTCGTGGCTTTTAAAGTTCATGAGCG 5953
DB 5885 TCTGGCGGCTTATGGAGAGGAGTGGCGGCGGCTCGTGGCTTTTAAAGTTCATGAGCG 5944
QY 5954 CGAGGTGCCCTCCACCGAGGACCTGGTCAACTTACTCTCCGTCATCTCTCTCTGGTGC 6013
DB 5945 CGAGATGCCCTCCACCGAGGACCTGGTCAACTTACTCTCTGCCATCTCTCTCTCTGGC 6004
QY 6014 CTTGGTCTGGGCTGTGTGGAGCAATATCTGGTGGGAGTGGGCGCGGAGAGG 6073
DB 6005 CTTGGTCTGGGCTGTGTGGAGCAATATCTGGTGGGAGTGGGCGCGGAGAGG 6064
QY 6074 GCGTGTGAGTGTGAACCGGCTGATAGGCTTCGCTTCGCGGGGTAAACCGTCTCCCG 6133
DB 6065 GCGTGTGAGTGTGAACCGGCTGATAGGCTTCGCTTCGCGGGGTAAATCATGTTTCCCG 6124
QY 6134 TACGCACTATGTGCTGAGAGGACGCTGCAGCAGTGTCACTCAGATCCTCTCTAGCCT 6193
DB 6125 CACGCACTATGTGCTGAGAGGACGCGCGAGCGGTGTACTCAGATCTCTCTCAGCCT 6184
QY 6194 TACCATCACTCACTGCTGAAGCGGCTCCACAGTGGATTATAGGAGTGTCTACGCG 6253
DB 6185 TACCATCACTCACTGCTGAAAAGGCTCCACAGTGGATTATAGGAGTGTCTCACAC 6244
QY 6254 ATGCTCCGCTCGTGGCTAAGGATGTTTGGGATTGGATATGCACGGTGTGACTGACTT 6313

DB 6245 GTGTTCCGGCTCGTGGCTAAAGGATGTTTGGGATGATATGCAAGGCTGTGACTGACTT 6304
QY 6314 CAAGACCTTGGCTCCAGTCCAAATCTCTGCCGCGGTACCGGGAGTCCCTTTCCTGTCTAG 6373
DB 6305 CAAGACCTTGGCTCCAGTCCAAAGCTCTCTGCCGCGAGTACCTTGGAGTCCCTTTTCTCGT 6364
QY 6374 CCAACGCGGTACAAAGGAGTCTGGCGGGGAGCGGATCATGCAAAACCACTCGCCCATG 6433
DB 6365 CCAACGCGGTACAAAGGAGTCTGGCGGGAGACGGGATCATGCAAAACCACTCGCCCATG 6424
QY 6434 CGGAGCACAGATCCCGCGGACATGTCAAAAACGGTTCCATGAGGATCGTAGGCCCTAGAAC 6493
DB 6425 TGGAGCACAGATCACCGGACATGTCAAAAACGGTTCCATGAGGATCGTGGGCTTAAGAC 6484
QY 6494 CTGAGCAACACGTTGGGACGGAACTTCCCATCAACGCATACACGAGGACCTTGAC 6553
DB 6485 CTGAGCAACACGTTGGGATGGAACTTCCCATCAACGCATACACGAGGCTTCGAC 6544
QY 6554 ACCCTCCCGCGGCCCAACTATTCCAGGGCTATGCGGGTGTGCTGAGGAGTACGT 6613
DB 6545 ACCCTCTCCAGCGCAAACTATTCTAGGGCTGTGGCGGTGGCGCTGAGGAGTACGT 6604
QY 6614 GGAGTTACGCTGTGGGGATTTCCACTAGTGAACGGGATGACCACTGAACGTA 6673
DB 6605 GGAGTTACGCTGTGGGGATTTCCACTAGTGAACGGGATGACCACTGAACGTA 6664
QY 6674 GTGCCATGCGAGTTCCGGCCCCGGAATTTCTTACGGAGTGGATGGAGTGGCTTGA 6733
DB 6665 GTGCCATGCGAGTTCCGGCTCTGAAATTTCTTCCGAGGTGACGAGTGGCTTGA 6724
QY 6734 CAGGTACGCTCCGGGTGCAACCTTCTTACGGAGGACGTCAGTTCCAGGCTCGGCT 6793
DB 6725 CAGGTACGCTCCGGGTGAGGCTTCTTACGGAGGAGTTTACATTCAGGTTGGCT 6784
QY 6794 CAACCAATACCTTGTTCGGGTGCGAGCTCCCATGCGAGGCCGAAACGGAAGTGC 6853
DB 6785 CAACCAATACCTTGTTCGGGTGCGAGCTACCATGCGAGGCCGAAACGGAAGTGC 6844
QY 6854 TACTTCATGCTACCGATTCCTCCCATATACAGAGAGAGCGGTAAGCGTAGGCTGGC 6913
DB 6845 CACTTCCATGCTACCGACCTCCCATCATCACAGCAACGCTAAGCGTAGGTTGGC 6904
QY 6914 TAGAGGCTTCCCGCTCTTTAGCCAGCTCATAGCTAGCAGTTGCTCGGCTTCCTT 6973
DB 6905 CAGGGGTCTCCCCCTCTTTGGCCAGCTCTTACAGTACGAGTTGCTGCGCTTCCT 6964
QY 6974 GAAGCGACATGCACTACCCACCATGACTCCCCCGGACGCTGACCTCATCGAGGCCAACCT 7033
DB 6965 GAAGCGACATGCACTACCCACCATGCTCTTCGAGACGCTGACCTCATCGAGGCCAACCT 7024
QY 7034 CTTGTGGCGGAGAGATGGCGGAAACATCATCTCGGTGAGTGCAGAGATAGGTTAGT 7093
DB 7025 CTTGTGGCGGAGAGATGGCGGAAACATCACCGCTGGAGTGCAGAGAACAGTGGT 7084
QY 7094 AATTCTGAGCTTTCGAAACGCTTACCGCGAGGGGATGAGAGGAGATATCGTCCG 7153
DB 7085 AGTCTGAGCTTCTTCGACCCGCTTCGAGCGGAGGAGTGCAGGAGTATTCGTTCC 7144
QY 7154 GCGGAGATCTTCGGAATAATCCAGAGTTCCCTCAGCTTGGCCCATATGGGACGCCC 7213
DB 7145 GCGGAGATCTTCGGAATAATCCAGAGTTCCCGCAGGATGCCCATCTGCGCGCGCC 7204
QY 7214 GGAATCAATCTCTCATGTGTAGTCTCGAAGGACCGGACATAGCTCCCTCCCGTGT 7273
DB 7205 GGAATCAATCTCTCATGTGTAGTCTCGAAGGACCGGACATAGCTCCCTCCCGTGT 7264
QY 7274 ACAGGATGCCATTTGCCACCTCAACAGGCTCTCCAATACCACTCCACGAGAGAG 7333
DB 7265 GCACGGGTGCCGTTGCCACCTATACAGGCTCTCCAATACCACTCCACGAGAGAG 7324
QY 7334 GACGTTCTCTGACAGAAATCCAATGTCTTCTGCTTGGCGGAGCTCGCCACTTAAGAC 7393

7325	7325	GAGGTTGTCCTAA	CAGAGTCTCCGTGTCCTTCTGCTTAGCGGAGCTCGCTACTAAGAC	7384	
Qy	7394	CTTCGGTAGCTCCGGATCGT	CGGCGGTGATAGCGGCACCGGACCGCCCTTCTGACCT	7453	
	7385	CTTCGGCAGCTCCGAATCAT	CGGCGGTGACAGCGGCACGGCACCGCCCTTCTTGACCA	7444	
	7454	GGCCTCGGACGAGGTGACAA	AGGATCGAGCTTGAGTCGTACTCCTCCATCGCCCCCT	7513	
Db	7445	GGCCTCGGACGAGGTGACAA	AGGATCGAGCTTGAGTCGTACTCCTCCATCGCCCCCT	7504	
Qy	7514	TGAAGGGGAGCGGGGACCC	CGATCTCAGCGACGCGGTCTTGCTCTACCGTGAGTGAGGA	7573	
Db	7505	TGAGGGGGAAACCGGGGAC	CCCGATCTCAGTGACGGGTCTTGCTCTACCGTGAGTGAGGA	7564	
Qy	7574	GGCTAGTGAGGATGTCTGT	CTGTCTCAATGTCTCTATACGTGACAGGCGCCCTGATCAC	7633	
Db	7565	AGCTAGTGAGGATGTCTGT	CTGTCTCAATGTCTCTATACGTGACAGGCGCCCTGATCAC	7624	
Qy	7634	GCCATGGCTGCGGAGGAAG	TAAAGTCTGCCATCAACCCGTTCAGCAACTCTTTGCTGCTGG	7693	
Db	7625	GCCATGGCTGCGGAGGAAG	TAAAGTCTGCCATCAACCCGTTCAGCAACTCTTTGCTGCTGG	7684	
Qy	7694	TCACCAACATCGTCTTAC	CCCAACATCCCGCAGCGCAAGCTCCGGCAGAGAAGGT	7753	
Db	7685	CCACCAATCATGTTTATG	CCCAACATCTCGAGCGCAGGCTCGGCGAGAAGGT	7744	
Qy	7754	CACCTTTGACAGATGTG	CAAGTCTCTGGATGATCAATACC	CGGACGTACTCAAGGAGATGAA	7813
Db	7745	CACCTTTGACAGACTG	CAAGTCTCTGGACGACCACTACC	CGGACGTACTCAAGGAGATGAA	7804
Qy	7814	GGCGAAGCGTCCACAGTT	AAGGCTTAAGCTTCTATCTATAGAGAGGCGCTGCAAGCTGAC	7873	
Db	7805	GGCGAAGCGCTCCACAGTT	AAGGCTTAAGCTTCTATCTATAGAGAGGCGCTGCAAGCTGAC	7864	
Qy	7874	GCCCCACATTCGSCCAAA	TCCAAATTTTGGCTATGGGGCAAGAGCTCGCGAACCTATC	7933	
Db	7865	GCCCCACATTCGSCCAAA	TCCAAATTTTGGCTATGGGGCAAGAGCTCGCGAACCTATC	7924	
Qy	7934	CAGCAGGCGCGTTAA	CCACATCCGCTCCGTGCGGAGGACTTGTGGAAGACACTGAAAC	7993	
Db	7925	CAGCAAGGCGTTAA	CCACATCCGCTCCGTGCGGAGGACTTGTGGAAGACACTGTAAC	7984	
Qy	7994	ACCAATTGACACCAATCAT	TGGCAAAAGTAGGAGTTTCTGCGTCCAAACGAGAGAGG	8053	
Db	7985	ACCAATTGACACCAATCAT	TGGCAAAAGTAGGAGTTTCTGCTGCCAACGAGAGAGG	8044	
Qy	8054	AGGCGCAAGCCAGCTCG	CTTATCGTATCCAGAGCTCGGAGTTCGTGTATGCGAGAA	8113	
Db	8045	AGGCGCAAGCCAGCTCG	CTTATCGTATCCAGAGTCTGGAGTTCGTGTATGCGAGAA	8104	
Qy	8114	GATGGCCCTTTACGACGT	GTCTCCACCTTCCTCAGGCGGTGATGGCTCCTCATACGG	8173	
Db	8105	GATGGCCCTTTACGACGT	GTCTCCACCTTCCTCAGGCGGTGATGGCTCCTCATACGG	8164	
Qy	8174	ATTTCAATPACTCC	CCCAAGCAGCGGTGAGTTCTCGGTGAATACCTGGAAATCAAAGAA	8233	
Db	8165	ATTTCAAGTACTCT	CTCGGCGCAGCAGTCCAGTTCTCGTGAATACCTGGAAATCAAAGAA	8224	
Qy	8234	ATGCCCTATGGCTCTCAT	ATGACACCCGCTGTTTGACTCAACGCTCACTGAGAGTGA	8293	
Db	8225	AAACCCCATGGCTTTT	CATATGACACTCGCTGTTTGACTCAACGCTCAACGAGAACGA	8284	
Qy	8294	CATTCGTGTGAGGAGT	CAATTTACCAATGTTGTGACTTTGGCCCCGAGGCGACAGAGC	8353	
Db	8285	CATTCGTGTGAGGAGT	CAATTTACCAATGTTGTGACTTTGGCCCCGAGGCGACAGAGC	8344	
Qy	8354	CATAAGTTCGTCA	CAGAGCGCTTTTACATCGGGGTCCTCCTGACTTAATCTCAAAGGCGCA	8413	
Db	8345	CATAAATCGCTCA	CAGAGCGCTTTTATATCGGGGTCCTCCTGACTTAATCTCAAAGGCGCA	8404	
Qy	8414	GAACTGCGGTTAT	TCGCGGTCGCGGAAGTGGCGTGCTGACGACTAGCTCGGTATATAC	8473	
Db	8405	GAACTGCGGTTAT	TCGCGGTCGCGCGAGGCGGTGCTGACGACTAGCTCGGTATATAC	8464	

QY	8474	CCTC	CATGTTACTTGAAGGCCACTCGAGCCTCTCGAGCTGCAAAAGCTCCAGACTGCGAC	8533
DB	8465	CCTC	CATGTTACTTGAAGGCCCTCTCGAGCCTCTCGAGCTGCGAGCTCCAGACTGCGAC	8524
QY	8534	GATG	CTCGTGAACGGAGACGACCTTGTGTTATCTGTGAAGCGCGGGAAACCGAGAGGA	8593
DB	8525	GATG	CTCGTGAACGGAGACGACCTCGTCTGTTATCTGTGAAGCGCGGGAAACCGAGAGGA	8584
QY	8594	TGGG	CGGCGCCCTACGAGCCTTTCACGAGGCTATGACTTAGGTATTCGCGCCCCCGCGGGA	8653
DB	8585	CGCG	CGAGCCTACGAGCTTCTTCACGGAGGCTATGACTTAGGTATTCGCGCCCCCGCGGGA	8644
QY	8654	TCCG	CCCCCAACCAAGATACGACCTGGAGCTGATAAATCATGTTCTCTCCAATGTGTCAGT	8713
DB	8645	CCCC	CCCCCAACCAAGATACGACTTGGAGCTGATAAATCATGTTCTCTCCAATGTGTCTGGT	8704
QY	8714	CGCG	CAGATGCATCTGGCAAAAGGTPATACTACCTCACCCGTGACCCCAACACCCCCCT	8773
DB	8705	CGCC	CAGATGCATCAGGCAAAAGGTTGTAACCTCACCCGTGATCCCAACACCCCCCT	8764
QY	8774	TGCA	CGGCTGCGTGGGAGACAGCTGAGACACATCCCAATCAACTCTTGGCTAGGCATAT	8833
DB	8765	AGCA	CGGCTGCGTGGGAGACAGCTGAGACACATCCCAATCAACTCTTGGCTAGGCACAT	8824
QY	8834	CATC	ATGTATGCGGCCACCCCTATGGGCAAGATGATTCGATGACTCACATTTTCTCTCCAT	8893
DB	8825	TATT	ATGTATGCGGCCCACTTTGTGGGCAAGATGATTCGATGACTCACATTTCTCTCCAT	8884
QY	8894	CTTT	CTAGCTCAAGAGCAACTTGAAGAGCCCTGGATTGTGAGATCTACGGGGCTTGCTA	8953
DB	8885	CTTT	CTAGCGCAGAGGCACTTGAAGAGCCCTGACTCCAGATCTACGGGGCTTGTTA	8944
QY	8954	CTCC	ATTGAGCCAATTGACCTTACTCTAGATCAATTGAACGACTCCATGGTCTTAGCGCATT	9013
DB	8945	CTCC	ATTGAGCCAATTGACCTTACTCTAGATCAATTGAACGACTCCATGGCCTTAGCGCATT	9004
QY	9014	TACAC	TCAAGTTACTCTCCAGGTGAGATCAATAGGGTGCTTCATGCCCTCAGGAACT	9073
DB	9005	TTCAC	TCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGCTTCATGCCCTCAGGAACT	9064
QY	9074	TGGGG	TACACCTTTGCGAACTCGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACT	9133
DB	9065	TGGGG	TACACCTTTGCGAGTCTGGAGACATCGGGCCAGGAGCGTCCGCGCTAGGCTACT	9124
QY	9134	GTCC	CAGGGGGGAGGGCGCCACTTGTGGCAGATACCTCTTTAACTGGGGCAGTAAGGAC	9193
DB	9125	GTCC	CAGGGGGGAGGGCGCCACTTGTGGCAAAATACCTCTTCAACTGGGCAAGTAAGGAC	9184
QY	9194	CAAG	CTTTAACTCACTCAATCCCGGCGCGCTCCAGCTGCACTTGTCTGGCTGGTTCGT	9253
DB	9185	CAAA	CTTTAACTCACTCAATCCCGGCTCGCTCCCGGCTGCACTTGTCTGGCTGGTTCGT	9244
QY	9254	CGCT	GGTTACAGCGGGGAGACATATACAGCCTGTCTCGTGGCCGACCCCGCTGGTT	9313
DB	9245	TGCT	GGTTACAGCGGGGAGACATATACAGCCTGTCTCGTGGCCGACCCCGCTGGTT	9304
QY	9314	TCCG	TGTGCCCTACTCTCTACTTTCTGTAGGGGTAGGCAATTTACCTGTCCCCAACCCGATG	9373
DB	9305	CATG	CTGTGCCCTACTCTCTACTTTCTGTAGGGGTAGGCAATTTACCTGTCCCCAACCCGATG	9364
QY	9374	AACG	GGAGCTAACCACTCCAGGCTTTAAGCCATTTCTGTGTGTGTGTGTGTGTGT	9425
DB	9365	AACG	GGAGATAAACACTCCAGGCAATAGGCAATCCCGCTTTGTGTGTGTGTGTGTGT	9416

RESULT 7
US-08-823-895A-27
; Sequence 27, Application US/08823895A
; Patent No. 6433159
; GENERAL INFORMATION:
; APPLICANT: Kevin P. Anderson
; TITLE OF INVENTION: Compositions And Methods For

TITLE OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 66 E. Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,895A
FILING DATE: March 17, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/453,085
FILING DATE: May 30, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/945,289
FILING DATE: September 10, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 810-1454
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 9416
TYPE: Nucleic
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO

Query Match	85.2%;	Score 8170.4;	DB 3;	Length 9416;
Best Local Similarity	91.8%;	Pred. No. 0;		
Matches 8636;	Conservative	0;	Mismatches 776;	Indels 0; Gaps 0;
14	TGGGGCGGACACTCCACCATGAATCACTCCCTGTGAGGAACTACTGTCTTCACGCAGAA	73		
yy				
bb				
5	TGGGGCGGACATCCACCATAGATCACTCCCTGTGAGGAACTACTGTCTTCACGCAGAA	64		
yy				
bb				
74	AGCGTCTAGCCATGGCGTTAGTATAGAGTGTGTCAGAGCCTCCAGAACCCCCCTTCCCGGG	133		
yy				
bb				
65	AGCGTCTAGCCATGGCGTTAGTATAGTGTGCTGTGAGCCTCCAGAACCCCCCTTCCCGGG	124		
yy				
bb				
134	AGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAAATTGCCAGAGCACCGGGTCC	193		
yy				
bb				
125	AGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAAATTGCCAGAGCACCGGGTCC	184		
yy				
bb				
194	TTTCTTTGGATCAACCCGCTCAATGTGCTGGAGATTTTGGCGTGCCCCCGCGAGACTGCTAG	253		
yy				
bb				
185	TTTCTTTGGATCAACCCGCTCAATGTGCTGGAGATTTTGGCGTGCCCCCGCGAGACTGCTAG	244		
yy				
bb				
254	CCGAGTAGTGTGGGTGCGGAAAGGCCTTGTGGTACTGCTGATAGGGTGTCTGCGAGTG	313		
yy				
bb				
245	CCGAGTAGTGTGGGTGCGGAAAGGCCTTGTGGTACTGCTGATAGGGTGTCTGCGAGTG	304		
yy				
bb				
314	CCCCGGGAGGTCTCGTAGACCGGTGCACCATGACGACGAATCCTAAACCTCAAAGAAAAAC	373		
yy				
bb				
305	CCCCGGGAGGTCTCGTAGACCGGTGCACCATGACGACGAATCCTAAACCTCAAAGAAAAAC	364		
yy				
bb				
374	CAAACGTAAACCAACCGCGGCCACAGAGAGTCAAGTTCGCGGGCGGTGTGCATCGT	433		
yy				
bb				
365	CAAACGTAAACCAACCGCGGCCACAGAGAGTCAAGTTCGCGGGCGGTGTGCATCGT	424		
yy				
bb				
434	TGCTGGAGTTTACTCTGTTCGCGCGCAGGGGCCCCCAGGTTGGGTGTGTCGCGGCATAGAA	493		
yy				
bb				

425	Db		TGTTGAGTTTACCTTTCCCGCGCAGGGGCCCCAGGTTGGGTGTGCGCGGCCACAGAA	484
494	Qy	GGCTTCCGAGCGGTGCAAACCTCGTGGAAAGGGGACAAACCTATCCCAAAGGCTCGCCGACC	553	
485	Db	GACTTCCGAGCGGTGCACCACTCGTGGAAAGGCGACAACTATCCCAAAGGCTCGCCGCCC	544	
554	Qy	CGAGGCGAGGGCGTGGGCTCAGCCGGGTACCTTGGCCCTCTCTATGGCAATAGAGGCCT	613	
545	Db	CGAGGCGAGGACCTGGGCTCAGCCGGGTACCTTGGCCCTCTCTATGGCAATAGAGGCCT	604	
614	Qy	GGGGTGGGAGGATGGCTCCTGTCAACCCGCGGCTCCCGGCTAGTTGGGGCCCCACGGA	673	
605	Db	AGGGTGGGAGGATGGCTCCTGTCAACCCGCGGCTCCCGGCTAGTTGGGGCCCCACGGA	664	
674	Qy	CCCCCGGCTAGGTCCGCTAACTTTGGGTAAAGGTCAATCGATACCCCTTACATGCGGCTTCGC	733	
665	Db	CCCCCGGCTAGGTCCGCTAACTTTGGGTAAAGGTCAATCGATACCCCTCAGTCGGCTTCGC	724	
734	Qy	CGATCTCATGGGTACATTTCCGCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTT	793	
725	Db	CGATCTCATGGGTACATTTCCGCTCGTCCGCGCCCCCTAGGGGGCGCTGCCAGGGCCCTT	784	
794	Qy	GGCAGACGGTGTCCGGGTTCTGGAGAGCGGCGTGAATATGCAACAGGAGNACTTGCCTCG	853	
785	Db	GGCAGACGGTGTCCGGGTTCTGGAGAGCGGCGTGAATATGCAACAGGAGNACTTGCCTCG	844	
854	Qy	TTGCTCTTTTCTCATCTTCTCTTTGGCTCTGCTGTCTCTGTTTGCACCATCCAGCTTCCGC	913	
845	Db	TTGCTCTTTTCTCATCTTCTCTTTGGCTCTGCTGTCTCTGTTTGCACCATCCAGCTTCCGC	904	
914	Qy	TTATGAAGTGCAGCAACGCTGTCCGGATATACCATGTACAGAACGACTGTCTCAACTCAAG	973	
905	Db	TTACGAAGTGCACAACGCTGTCCGGATATATCATGTACAGAACGACTGTCTCAACGCAAG	964	
974	Qy	CATTGTGTATGAGGACGCGACGTGATCATGCACTCTCCGGTGGTGGCCCTGTGTTCA	1033	
965	Db	CATTGTGTATGAGGACGCGACTTGATCATGCACTCTCTGGTGGTGGCCCTGCGTTCG	1024	
1034	Qy	GGAGGTTACAGCTCCGCTTGTGGGTAGCGCTCACTCCACGCTCGCGCCACAGGAATGC	1093	
1025	Db	GGAGGCAACTCTCCGCTGTGGGTAGCGCTCACTCCACGCTCGACCCAGGAACGT	1084	
1094	Qy	CAGCGTCCCCACTACGACAATACAGCCACGCTCGACTTGTCTCGTTGGGACGGCTGCTTT	1153	
1085	Db	CACCATCCCCACACAGCAGATACGACCCACGTCGATCTGCTCGTTGGGCGGCTGCTTT	1144	
1154	Qy	CTGCTCGCTATGTACGTGGGGGATCTGCGGATCTATTTTCTGTCTCCGACGTGTT	1213	
1145	Db	CTGTTCGCTATGTACGTGGGGGACCTCTCGGATCTGTTTCTGTCTCTCAGCTGTT	1204	
1214	Qy	CACCTTCTCGCTCGCGGATCAGACAGTGCAGGACTGCACTGTCTCAATCTATCCCGG	1273	
1205	Db	CACCTTCTCGCTCGCGGATGTGACATTTACAGGACTGTAACTGCTCAATTTATCCCGG	1264	
1274	Qy	CCATGTATCAGGTCACCGCATGGCTTGGGATATGATGAATGTTGCTACCTCAACAGC	1333	
1265	Db	CCATGTTCGGGTCAACGATATGCTTGGGACATGATGAATGTTGCTCGCCCAACAGC	1324	
1334	Qy	CCTAGTGGTGTCCAGTTGTCTCCGATCCCAACGCTGTGTTGACATGTTGGCGGGGGC	1393	
1325	Db	CCTAGTGGTGTCCAGTTACTCCGGATCCCAACGCTGTGTTGACATGTTGGCGGGGGC	1384	
1394	Qy	CCACTGGGAGTCTTGGCGGGCTTGCCTACTATTCATGGTAGGAACTGGGGCTAAGGT	1453	
1385	Db	CCACTGGGAGTCTTGGCGGGCTTGCCTACTATTCATGGCGGAACTGGGGCTAAGGT	1444	
1454	Qy	TCTGATTGTGGCGCTACTCTTTTGGCGGCTTGACGGGGAGACCCACACGACGGGGAGGT	1513	
1445	Db	TCTGATTGTGATGCTACTTTTGTGCGCTTGACGGGGATACCCACGTGACAGGGGGGC	1504	
1514	Qy	GGCGGGCCACCAACCTTCGGGTTACAGTCCCTTTTCTATCTGGGGCGCTCTAGAAAT	1573	

Db 1505 GCAAGCAAAACCAACACAGGCTCGTGTCAATGTTGCAAGTGGCCGCTCTCAGAAAT 1564
Qy 1574 CCAGCTTGTGAATACCAACGCGAGCTGGCACATCAACAGGACTGGCCCTAAATGCAATGA 1633
Db 1565 CCAGCTTATAAACCAACNATGGGAGTTGGCACATCAACAGGACTGGCCCTGAACTGCAATGA 1624
Qy 1634 CTCCCTCCAAACTGGGTTCTTTGGCCGCTGTTTTCACGACACAAGTTTCAACTGTCGCG 1693
Db 1625 CTCTCTCCAGACTGGGTTCTTTGGCCGCTGTTTACACACATAGTTTCAACTCTGTCGCG 1684
Qy 1694 GTGCCCGAGCGCATGCCAGCTGCCGCCCATTTGACTGTTGCGCCACGAGGTTGGGGCC 1753
Db 1685 GTGCCAGAGCGCATGGCCAGTGGCCGACCATTTGACAAAGTTTCGACAGGATGGGGTCC 1744
Qy 1754 CATCACTATACCTAACCTAACAGCTCGGATCAGAGGCTTATTTGCTGGCATTTACGCGCC 1813
Db 1745 CATTACTTATGCTGAGTCTAGCAGATCAGACAGAGGCCATATTTGCTGGCACTACCCACC 1804
Qy 1814 TCGACCGTGTGTGTCGTATCCCGCGTGCAGGTGTGTGTCCAGTGTATTTGTTTCACCCC 1873
Db 1805 TCCACAATGTACCATCGTACCTCGTGGAGGTGTGGGCCAGTGTACTGCTTCACCCC 1864
Qy 1874 AAGCCCTGTGTGGTGGGACCAACGATCGTTCGGTGTCCCTACGTATAGCTGGGGGA 1933
Db 1865 AAGCCCTGTGTGGTGGGACGACGATCGTTCGGTGTCCCTACGTATAGATGGGGGA 1924
Qy 1934 GAATGACAGAGTGTGTCCTCAACACAGCGTCCGCCAACAGGCAACTGGTTCGG 1993
Db 1925 GAAACGAGACTGACGTGTCTCAACAAACACGCGGCCGCCGAGGAACTGGTTCGG 1984
Qy 1994 CTGTACATGATGAATAGTACTGGGTTCACTAAGACGTGGGAGGTCCCCCGTGTAAACAT 2053
Db 1985 CTGCACATGATGAATAGCACCGGTTCAACAGACATGTGGGGGCCCCCGTGTAAACAT 2044
Qy 2054 CGGGGGGTTCGTAACCGCACTTTGATCTGCGCCCAACGCACTGTTCCGGAGAACCCCGA 2113
Db 2045 CGGGGGGTTCGGCAACCAACACCTGACTGCCCCACGCACTGTTTCCGGAAGCACCCCGA 2104
Qy 2114 GGCTACTTACAAAAATGTGGCTCGGGGCTGTTGACACCTAGGTGCTAGTACACTA 2173
Db 2105 GGCTACTTACAAAAATGTGGTTCGGGGCTTGGCTGACACCTAGGTGCTAGTGTGACTA 2164
Qy 2174 CCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTTCATCTTTAAGGTTAGGAT 2233
Db 2165 TCCATACAGGCTTCGGCACTTACCCCTGCACTGTGTAACTTTACCATCTTCAAGTTAGGAT 2224
Qy 2234 GTATGTGGGGGCGTGGAGCACAGGCTCAATGCCGCATGCAATTTGGAATCTGAGAGAGCG 2293
Db 2225 GTATGTGGGGGCGTGGAGCACAGGCTCAATGTCTGCAATTTGGACCCCGAGGAGAGCG 2284
Qy 2294 CTGTAACTTGGAGACAGGGATAGGTGAGAACTCAGCCCGTGTGCTGTCTACACAGA 2353
Db 2285 TTGTGACTTGGAGACAGGGATAGGCCGAGGCTCAGCCCGTGTGCTGTCTACACAGA 2344
Qy 2354 GTGGCAGATAGTCCCTGTCTTACCAACCTACCGCTTTATCCACTGTTTGTGATCCA 2413
Db 2345 GTGGCAGATAGTGGCCCTGTTCTTACCAACCTACAGCTGTCTCAGTGGCTGTGATCCA 2404
Qy 2414 TCTCCATCAGAACATCGTGGAGTGCATATCCTGTACGGTGTAGGGTCAGCGTTTGTCTC 2473
Db 2405 CCTCCATCAGAACATCGTGGAGTGCATATCTATACGGTATAGGGTCAGCGTTTGTCTC 2464
Qy 2474 CTTTGCATCAATGGGAGTACATCCTGTGTGTTTTTCTTCTCTGGCAGAGCGCGGCT 2533
Db 2465 CTTTGCATCAATGGGAGTGTCTCTGTGTTTTTCTTCTCTCTAGCGGACGACGCTGT 2524
Qy 2534 GTGTCCCTGTGTGGATGATGCTCTGTATAGCCAGGCTGAGGCGGCTTAGAGAACTT 2593
Db 2525 CTGTCCCTGTGTGGATGATGCTCTGTATAGCCAGGCGGAGGCGGCTTGGAGAACT 2584
Qy 2594 GGTGTCTCTCAATGGGGGCTCGTGGCGGAGCGATGATTTCTCTCTCTTCTTGTGT 2653
Db 2585 GGTGTCTCTCAATCGGGCTGTGTGGCGGCGCACATGGCATCTCTCTCTCTTCTTGTGT 2644

Qy 2654 CTTTGTGGCCGCTGGTACATTAAAGGCGAGGCTGCTCTCTGGGCGGCGTATGCTTTTTA 2713
Db 2645 CTTTGTGGCCGCTGGTACATTAAAGGCGAGGCTGCTCTCTGGGCGGCAATATGCTCTTTA 2704
Qy 2714 TGGCGTATGGCGCTGCT 2773
Db 2705 TGGCGTATGGCGCTGCT 2764
Qy 2774 CCGGAGATGGTGCATCGTGGGGGGTGGGTTCTTGTAGTCTGGTATTTCTTGACCTT 2833
Db 2765 CCGGAGATGGTGCATCGTGGAGGGCGGTTTTTGTGGGTCTGGTACTCTGACTTT 2824
Qy 2834 GTCAACATACATAAAGTGTCTCACTAGGCTCATATGCTGTGTTTACAATATCTTTATCAC 2893
Db 2825 GTCAACATACATAAAGTGTCTCTCGTAGGCTCATATGCTGTGTTTACAATATTTTACCAC 2884
Qy 2894 CAGAGCGAGGCGCACATGCAAGTGTGGGTCCCCCCCCCTCAACGTTTCGGGAGGCGCGA 2953
Db 2885 CAGAGCGAGGCGGACTTACATGTGTGGATCCCCCCCCCTCAACGCTCGGGAGGCGCGA 2944
Qy 2954 TGCATATCT 3013
Db 2945 TGCATATCT 3004
Qy 3014 CTTGCTCGCATCTCTCGGCCGCTCATGTGCTCCAGGCTGGCATAAACGAGAGTGGCGTA 3073
Db 3005 TCTAATGGCATCTCTCGTCCGCTCATGTGTCTCAAGCTGGCATAAACGAGTGGCGTA 3064
Qy 3074 CTTGCTCGGCTCAAGGCTCATCTGTGCTCATGTGTTAGTCCGAAAGTCCGCCGGGG 3133
Db 3065 CTTGCTCGGCTCAAGGCTCATCTGTGCTCATGTGCTCAAGTGGCATAAACGAGTGGCGTA 3124
Qy 3134 TCATTATGTCCAAATCGTCTTCAATGAAGCTGGGCGGCTGACAGGTAAGTACGTTTATAA 3193
Db 3125 TCATTATGTCCAAATGGCTTTCATGAAGCTGGGCGGCTGACAGGCAAGTACATTTACAA 3184
Qy 3194 CCATCTTACCCCACTCTGGGCACTGGGCGCACGGGGCTTACGAGACTTTCGGTGGCGGT 3253
Db 3185 CCATCTTACCCCGCTACTGGGATTTGGCCACGCGCGGGCTTACGAGACTTTCGGTGGCGGT 3244
Qy 3254 AGAGCCGCTGCTCTTCCGCAATGGAGACCAAGGTCATCACCTGGGGAGCAGACACCGC 3313
Db 3245 GAGCCGCTGCTCTTCCGCAATGGAGACCAAGATCATCACCTGGGGAGCAGACACCGC 3304
Qy 3314 TGGTGTGGGCACTCATCTTGGGTCTACCCGCTCTCCGCGCCGAGGGGAGGAGATATT 3373
Db 3305 GGGTGTGGGCACTCATCTTGGGTCTGCGCTCTCCGCGCGAGGGGAGGAGATACT 3364
Qy 3374 TTTGGGACCGGCTGATAGTCTCGAAAGGCAAGGTTGGGAGTCTCTTGGGCCCATCACGGC 3433
Db 3365 CTTGGGCCCGGCGATAGTCTTGAAGGCGGGGTTGCGACTCTCTCGCGCCCATCACGGC 3424
Qy 3434 CTACTCCCAACAAACGCGGGGCTACTTGGTGTGATCATCATCTAGCTCTCAGGCCGCGA 3493
Db 3425 CTACTCCCAACAGACGCGGGGCTACTTGGTGTGATCATCATCTAGCTCTCAGGCCGCGA 3484
Qy 3494 CAAGAACAGGTCGAAGGGAGGTTCAAGTGGTGTCTACCGCAACACAACTTTTCTCTGGC 3553
Db 3485 CAAGAACAGGTCGAGGAGAGGTTTCAAGTGGTGTCTACCGCAACACAACTTTTCTCTGGC 3544
Qy 3554 GACCTGCATCAACCGGCTGTGCTGGACTGTCTACCTAGGCGCTGGCTCGAAGACCTTAGC 3613
Db 3545 GACCTGCATCAACCGGCTGTGCTGGACTGTGCGCTTACCATGGTGTGGCTCAAGAGCCTTAGC 3604
Qy 3614 CGGTCCAAAGGTCGAATCACCCAAATGTACCAAAATGTAGACCTGAGCTCTGCGGCTG 3673
Db 3605 CGGCCAAAGGGGCAATCACCCAGATGTACCTAAATGTGAGACGAGGACCTCTGCGGCTG 3664
Qy 3674 GCAGGCGCCCCCGGGGCGGCTCCATGACACCATGAGCTGTGAGCTCGGACCTTTA 3733
Db 3665 GCGGCGCCCCCGGGGCGGCTTCTTGACACCATGACCTGTGGCAGCTCAGACCTTTA 3724

Qy	3734	CTTGGTCA	CAGAGACA	TGCTGAT	GTATTC	CGGTG	CGCGCGG	AGGCGA	CAGCAGGGG	AAG	3799
Db	3725	CTTGGTCA	CAGAGACA	TGCTGAT	GTATTC	CGGTG	CGCGCGG	AGGCGA	CAGTATAGGGG	GAG	3784
Qy	3794	TCCTACT	CTCTCC	CAGGCCG	CTCTCC	TACTT	GAAGG	CTCTCC	GGGTGGT	CCATTTGCTTTG	3853
Db	3785	CTTGCTCT	CCCCCAGG	CGCTGTCT	CTACTT	TGAAGG	GTCTTT	CGGGTGGT	CCATCTGCTCTG		3844
Qy	3854	CCCTTCC	GGGCA	CGTCTG	GTGGCG	GTCTTT	CCGGGTG	CTGTGTG	CAACCGGGGGG	CTCGCAA	3913
Db	3845	CCCTTCC	GGGCA	CGCTGT	GGGCA	CTTTCC	GGGCTG	CGGTATG	CACCGGGGGT	TGGAA	3904
Qy	3914	GGCGGT	GGACTT	CATAC	CCGTTG	AGTCTA	TGGA	AACTPAC	CATCGG	TCTCCGTTCTTAC	3973
Db	3905	GGCGGT	GGACTT	TGTG	CCCGTAG	AGTCCAT	TGGA	AACTTACT	TATCGG	TCTCCGTTCTTAC	3964
Qy	3974	AGACAA	CTCAAC	CCCCCGG	CTGTAC	CGCAGAC	ATTC	CAAGTGG	CAATCTG	CACGCTCC	4033
Db	3965	GGACAA	CTCTAT	CCCCCCCCG	CGTAC	CGCAGT	CAATTC	CAAGTGG	CCACCTCA	CACGCTCC	4024
Qy	4034	TACTGG	CAGCGG	CAAGAG	CAAC	AAAGTGC	CGGCTG	CGTATG	CAAGCCCA	AGGTACAAGGT	4093
Db	4025	CACTGG	CAGCGG	CAAGAG	TACTA	AGTGC	CGGCTG	CAATATG	CAAGCCCA	AGGTACAAGGT	4084
Qy	4094	GCTCGT	CCTGAA	CCCGTCC	GTGCG	CCCACTT	AGGGT	TTTGGG	CGTATATG	TCCAAGGC	4153
Db	4085	GCTCGT	CCTCAA	TCCGTC	CGTTG	CGCTACTT	AGGGT	TTTGGG	CGTATATG	TCTAAGGC	4144
Qy	4154	ACACGG	TATCGA	CCCTAAC	ATCAGAA	CTGGG	TAA	AGCACTT	ACACGGG	CGGCTCCAT	4213
Db	4145	ACACGG	TATTGAC	CCCCAAC	ATCAGAA	CTGGG	TAA	AGCACTT	ACCA	CAGGCGCCCCGT	4204
Qy	4214	TACGTAC	TCCACCT	ATGCGA	AGTTCCTT	GC	CAGCGG	TGGTGT	TCTGGG	GGGCGCTATGA	4273
Db	4205	CACATA	CTTACCT	TATGCGA	AGTTCCTT	GC	CAGCGG	TGGTGT	TCTGGG	GGGCGCTATGA	4264
Qy	4274	CATCAT	TAATATG	TGATGAG	TGCCA	CTCAA	CTGACT	CGACTAC	CAATCTT	GGGCACTCGGCAC	4333
Db	4265	CATCAT	TAATATG	TGATGAG	TGCCA	CTCAA	CTGACT	CGACTAC	CAATCTT	GGGCACTCGGCAC	4324
Qy	4334	AGTCTT	GGACCA	AGCGGAG	CGGTG	GAGCGG	CGCTCGT	GTGCTGCC	CAACCGGT	ACACC	4393
Db	4325	AGTCTT	GGACCA	AGCGGAG	CGGTG	GAGCGG	CGCTGT	GTGCTGCC	CAACCGGT	ACACC	4384
Qy	4394	TCCGGG	ATCGGTT	ACGTC	GCACAC	CCCCATAT	CGAGG	AAATAGG	CCGCTG	CCACCAATGG	4453
Db	4385	TCCGGG	ATCGGTT	ACCGT	GCACAC	CCCCAA	CAATCGAGG	AGTGGCC	CTGTCT	TAATCTGG	4444
Qy	4454	AGAGAT	CCCCCTT	CTATGG	CAAGG	CCATCCC	CAATGAG	CGCCATCA	AGGGGGG	AGGAGCATCT	4513
Db	4445	AGAGAT	CCCCCTT	CTATGG	CAAGG	CCATCCC	CAATGAG	CGCCATCA	AGGGGGG	AGGAGCATCT	4504
Qy	4514	CATTTT	CTGCCAT	TCCAA	GAGAA	AAATGTG	ACGACT	TCGCG	CAAGCTGA	CAGGCTTCGG	4573
Db	4505	CATTTT	CTGTCT	CAATCC	AAGAA	AGATG	CGAC	AGCTCG	CGCGCA	AGGTGTCA	4564
Qy	4574	ACTGAA	CGCTGT	AGCATAT	TACCGG	GGCCTTG	ATGTG	CTCGT	CAATAC	CGGCTATCGGAGA	4633
Db	4565	AATCAA	CGCTGT	GGCGTAT	TACCGG	GGGCTCG	ATGTG	CTCGT	CAATAC	CAATCTCGGAGA	4624
Qy	4634	CGTCTGT	CTGTGG	CAACAG	ACGCTCT	AA	TGACGGG	TTTACC	GGCGATTTT	TGACTCAGT	4693
Db	4625	CGTCTGT	CTGTGG	CAACAG	ACGCTCT	TA	TGACGGG	TTTACC	GGCGATTTT	TGACTCAGT	4684
Qy	4694	GATCGA	CTGCAAT	ATCATGT	GTCA	CCCGAC	AGCTCG	CACTT	CAGCTT	TGGATCCCACTTAC	4753
Db	4685	GATCGA	CTGTAA	CATGTGT	CA	CCCGAC	AGCTCG	CACTT	CAGCTT	TGGATCCCACTTAC	4744
Qy	4754	CATTGAG	CAGCAG	CCGTG	CCCCA	AGACCG	GGTTC	GGCTTC	GCGCA	ACGGGAGGTAGA	4813
Db	4745	CATTGAG	CAGCAG	CCGTG	CCCCA	AGCAG	CGTTC	GGCTTC	GCGCA	ACGGGAGGTAGA	4804
Qy	4814	TGGCAG	GGGTAG	GAGTGC	ATCTA	CAGG	TTTGTG	ACTT	CCAGG	AAACGGGCTCTCGGCAT	4873

4805	Db	 TGGCAGGGGTAGGAGAGGCATCTACAGGTTTGTGCTCGGGAGAACGGCCCTCGGGCAT	4864
4874	Qy	 GTTTCGATTCTTCGGTCTGTGTCAGTGCTATGACGGGGTGTGCTTGGTATGAGCTCAC	4933
4865	Db	 GTTTCGATTCTTCGGTCTGTGTCAGTGCTATGACGGGGTGTGCTTGGTACGAGCTCAC	4924
4934	Qy	 GCCCGTGAACCTCGGTTAGGTTGCGGGCTTACCTAAATACACCAAGGTTGCCGTCGTG	4993
4925	Db	 CCCGGCCGAGACCTCGGTTAGGTTGCGGGCTTACCTGMAACACACAGGGTTGCCCGTTTG	4984
4994	Qy	 CGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTCAGAGGCTCACCCACATAGATGCCCA	5053
4985	Db	 CCAGGACCACTCGGAGTTCTGGGAGAGTGCTTTCACAGGGCTCACCCATATAGATGCACA	5044
5054	Qy	 CTTCTGTGCCAGACTAAACAGGCAGGAGACAACTTTCTTACTGTGTGCATATCAAGC	5113
5045	Db	 CTTCTGTGCCAGACCAAGCAGGAGGAGACAACTTCCCTACTCTGTGTAGCATACCAAGC	5104
5114	Qy	 TACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAATGTGGAAGTGCT	5173
5105	Db	 CACGGTGTGCGCCAGGGCTCAGGCCCACTCCATCATGGGATCAAATGTGGAAGTGCT	5164
5174	Qy	 CATACGGCTGAACCTTACATCTGACAGGGGCCAAACACCCCTGCTGTATAGGCTAGGAGCCGT	5233
5165	Db	 CATACGGCTGAACCTTACGCTGACAGGGGCCAAACACCCCTGCTGTACAGGCTGGGAGCCGT	5224
5234	Qy	 CCAAATGAGGTCATCTCTCACACCCCAATAACTAAATACATCATGSCATGTCATGTCGGC	5293
5225	Db	 CCAAATGAGGTCATCTCTCACACCCCAATAACTAAATACATCATGSCATGTCATGTCGGC	5284
5294	Qy	 TGACCTGGAGGTCGTCACTAGACACCTGGGTGCTGGTAGGGCGAGTCTCTGCAGCTTTGGC	5353
5285	Db	 TGACCTGGAGGTCGTCACTAGACACCTGGGTGCTGGTAGGGCGAGTCTCTGCAGCTTTGGC	5344
5354	Qy	 CGCATACTGCTGAACAGCAGGAGTGTCATTGTGGGCAGGATCATCTTGTCCGGGAA	5413
5345	Db	 CGGCTATTGCTGACAAACAGGCAGTGTCATTGTGGGTAGGATTCATTGTCCGGAG	5404
5414	Qy	 GCCAGCTGTCTGTTCCGACAGGGAGTCTCTACACAGGAGTTCGATGAGATGGAGAGTG	5473
5405	Db	 GCCGGCCATTGTTCCCGACAGGGAGCTTCTCTACACAGGAGTTCGATGAAATGGAAAGAGTG	5464
5474	Qy	 TGCCCTCAAACTTCTTACATCGACAGGGAATGACAGCTGCCCGAGCAATTCAGACAAA	5533
5465	Db	 CGCTCGCACTTCCCTTACATCGAGCAGGGAATGACAGCTGCCCGAGCAATTCAGACAAA	5524
5534	Qy	 GGCGCTCGGTTGTGTGAAAACGGCCACCAAGCAAGCGAGGCTGCTCTCCCTGGTGGGA	5593
5525	Db	 AGCGCTCGGTTTACTGTCAACAGCCACCAACAAAGCGAGGCTGCTCTCCCTGGTGGGA	5584
5594	Qy	 GTCCAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCAATGTGGGAATTCATCAGCGG	5653
5585	Db	 GTCCAGTGGGAGCCCTTGAGACATTCCTGGGCGAAGCAATGTGGGAATTCATCAGCGG	5644
5654	Qy	 AATACAGTACTAGCAGGCTTATCCACTCTGCTCTGGAAACCCCGCATAGCATCATTTGAT	5713
5645	Db	 GATACAGTACTAGCAGGCTTATCCACTCTGCTCTGGAAACCCCGCATAGCATCATTTGAT	5704
5714	Qy	 GGCATTTACAGCTTCTATCATAGCCCGCTCACCCACCAAAACCCCTCTGTTTAAACAT	5773
5705	Db	 GGCATTCACAGCTTCTATCACAGCCCGCTCACCCACCAAAAGTACCCCTCTGTTTAAACAT	5764
5774	Qy	 CTTGGGGGATGGGTGGCTGCCMACTGCTCTCTCCAGCGCTGCGTACAGCTTTCGTGGG	5833
5765	Db	 CTTGGGGGGTGGGTGGCTGCCMACTGCTCTCTCCAGCGCTGCGTACAGCTTTCGTGGG	5824
5834	Qy	 CGCCGGCATCGCCGGTGGGTGGCTGTCGAGCATATAGGCTTGGGAAGGTGCTTGTGGACAT	5893
5825	Db	 CGCCGGCATCGCCGGTGGGTGGCTGTCGAGCATATAGGCTTGGGAAGGTGCTTGTGGACAT	5884
5894	Qy	 CTTGGCGGGCTATGGGCGAGGGGTAGCCGGCGCATCTGCTGGCCCTTTAAGTTCATGACGG	5953

5885 TCTGCGGGTTATGGAGCAGGAGTGCGCGCGCGCTTCTGTTGGCTTTAAAGTCAATGAGCGG 5944 Db
5954 CGAGGTGCCCTCCACCGAGACCTGTCAACTTACTCCCTGCCATCCTCTCTCTGCTGC 6013 QY
5945 CGAGATGCCCTCCACCGAGACCTGTCAACTTACTTCTTGCCATCCTCTCTCTGCGC 6004 Db
6014 CTTGTTCTGCGGGTCTGTGCGCAGCAATATCTGCGTCGCGCAGCTGGGCCCGGAGAGG 6073 QY
6005 CTTGTTCTGCGGGTCTGTGTGACCAATCTGCGTCGACACGTGGTCCGGGAGAGG 6064 Db
6074 GGTGTGCGAGTGAATGAACCGGCTGATAGCTTTCGCTTCCGGGGTAAACACGCTTCCCC 6133 QY
6065 GGTGTGCGAGTGAATGAACCGGCTGATAGCTTTCGCTTCCGGGGTAAATCATGTTCCTCC 6124 Db
6134 TAGCCTATGTCTGAGAGCAGCTGCGAGCAGCTGTCACTCAGATCCTCTCTAGCCT 6193 QY
6125 CAGCCTATGTCTGAGAGCAGCGCGCAGCGCTGTACTCAGATCCTCTCCAGCCT 6184 Db
6194 TACCATCACTCAACTGTGTAAGCGGTCCACAGTGGATTAATGAGGACTGTCTTAGGCC 6253 QY
6185 TACCATCACTCAGCTGCTGAAGAGGCTCCACAGTGGATTAATGAAGACTGTCTCCACCC 6244 Db
6254 ATCTCGGCTCGTGTAGGAGTGTTCGGATTTGGATGATGACCGTGTGACTGACTT 6313 QY
6245 GTGTTCCGGCTCGTGTAGGAGTGTTCGGACTGGATATGACCGTGTGTGACTGACTT 6304 Db
6314 CAAGACTGGCTCAGTCCAAACTCTCTGCGCGGTTTACCGGAGTCCCTTTCTCTGTCATG 6373 QY
6305 CAAGACTGGCTCAGTCCAAACTCTCTGCGCGGTTTACCGGAGTCCCTTTCTCTGTCATG 6364 Db
6374 CCAAACGCGGTTAAGGAGTCTGCGCGGCGGAGCGGATCATGCAACCAACCTGCGCCATG 6433 QY
6365 CCAAACGCGGTTAAGGAGTCTGCGCGGCGGAGCGGATCATGCAACCAACCTGCGCCATG 6424 Db
6434 CGAGACACATGCCCGGATGTCAAAACGTTTCCATGAGGATCGTAGGGCTAGAAC 6493 QY
6425 TGAGACACATGCCCGGATGTCAAAACGTTTCCATGAGGATCGTAGGGCTAGAAC 6484 Db
6494 CTCAGCAACAGTGGCAGCGAAGCTTCCCATCAACGCTACACACGCGACCTTGCCAC 6553 QY
6485 CTGAGCAACAGTGGATGGAACTTCCCATCAACGCTACACACGCGCCCTGCAC 6544 Db
6554 ACCCTCCCGCGGCCCAACTATTCCAGGGCGCTATGCGGGTGGCTGCTGAGGAGTACGT 6613 QY
6545 ACCCTCCCGCGGCCCAACTATTCTAGGGCGCTGTGCGGGTGGCTGAGGAGTACGT 6604 Db
6614 GGAGTTACGGTGTGGGGATTTCCACTAGTGAACCGGCAATGACCACTGACACGATAA 6673 QY
6605 GGAGTTACCGGGTGGGGATTTCCACTAGTGAACCGGCAATGACCACTGACACGATAA 6664 Db
6674 GTGCCATGCCAGTTCCGGCCCCCGAATCTTCCAGGAGTGGATGGAGTGGCGTTGCA 6733 QY
6665 GTGCCATGCCAGTTCCGGCTCTCTGAATCTTCTCGAGGTGGACGAGTGGCGTTGCA 6724 Db
6734 CAGTACGCTCCGGCTGCAAACTCTTCTACGGGAGGACGCTCACTGTTTCCAGTTCGGGCT 6793 QY
6725 CAGTACGCTCCGGCTGACGGCTCTCTTACGGGAGGTTACATTTCCAGTTCGGGCT 6784 Db
6794 CAAACCAATATTTGTTGGGTGCGAGTCCCATGCGAGCCGAAACGGACGTAAACAGTGTCT 6853 QY
6785 CAAACCAATATTTGTTGGGTGCGAGTACCATGCGAGCCGAAACGGATGTAGCAGTGTCT 6844 Db
6854 TACTTCCATGCTACCGATCCCTCCCATTTACAGCAGAGACGGCTAAGCGTAGGCTGGC 6913 QY
6845 CACTTCCATGCTACCGACCCCTCCCATCACAGCAGAAACGGCTAAGCGTAGGCTGGC 6904 Db
6914 TAGAGGTCTCCCGCTCTTTAGCCAGTCTCATCAGCTAGCCAGTTGTCTGCGCCCTCTTT 6973 QY
6905 CAGGGGTCTCCCGCTCTTTAGCCAGTCTCTTACGTAGCCAGTTGTCTGCGCCCTCTTT 6964 Db
6974 GAAGCGCATGCACTACCCACCATATGACTCCCGGAGCGTCACTCATGAGGCGCAACT 7033 QY
6965 GAAGCGCATGCACTACCCACCATATGACTCCCGGAGCGTCACTCATGAGGCGCAACT 7024 Db

7034 CTTTGTGGCGCAGGAGATGGGCGGAAACATCACTCCGTTGGAGTCAAGAAATAGGTAGT 7093 QY
7025 CTTTGTGGCGCAGGAGATGGGCGGAAACATCACTCCGCTGGAGTCCGAGAACAAAGTGT 7084 Db
7094 AATTCTGGAGTCTTTTCGAACCGCTTCAACGCGAGGGGATGAGAGGAGATATTCCTGCGC 7153 QY
7085 AGTCTGGAGTCTTTTCGACCCGCTTTCGAGGGAGGAGATGAGAGGAGATATTCCTGTC 7144 Db
7154 GCGGAGATCTCTGCGAAATCCAGGAATTCAGGTTCCCTCAGGTTGCCCATATATGGCAGCC 7213 QY
7145 GCGGAGATCTCTGCGAAATCCAGGAATTCAGGTTCCCTCAGGTTGCCCATATTCGGCGGCC 7204 Db
7214 GGAATCAATCTCTCACTGCTAGAGTCTTGGAGGACCGGACTACGTCCCTCCGTTGGT 7273 QY
7205 GGAATCAACCCCTCCACTGTTAGAGTCTTGGAGGACCGGACTACGTCCCTCCGTTGGT 7264 Db
7274 ACAGGATGCCATTCGCCACCTACCAAGCTCTCCAAATACCACTCCACGGAGAAAGAG 7333 QY
7265 GACGGTGGCGCTTCCACCTATCAAGGCCCTCCAAATACCACTCCACGGAGAAAGAG 7324 Db
7334 GACGGTGTCTGACAGAAATCCAAATGTCTTCTGCTTGGCGGAGTCCGCCACTAAGAC 7393 QY
7325 GACGGTGTCTTAAAGAGTCTTCTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 7384 Db
7394 CTTTGTGGTCCGATCGTCCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 7453 QY
7385 CTTTGTGGTCCGATCGTCCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 7444 Db
7454 GGCCTCCGAGCAGGTTGACAAAGATCCGAGTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 7513 QY
7445 GGCCTCCGAGCAGGTTGACAAAGATCCGAGTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 7504 Db
7514 TGAAGGGAGCGCGGGGACCCCGATCTCAGCGACGGTCTTGGTCTTACCGTGTAGTAGGA 7573 QY
7505 TGAAGGGAGCGCGGGGACCCCGATCTCAGCGACGGTCTTGGTCTTACCGTGTAGTAGGA 7564 Db
7574 GGTAGTAGGATGTCTGCTGCTCAATGTCTTATACGTGACAGCGCCCTTGATCAC 7633 QY
7565 AGCTAGTAGGATGTCTGCTGCTCAATGTCTTATACATGACAGCGCCCTTGATCAC 7624 Db
7634 GCATGGCTCGGAGGAAAGTAAAGTCCCATCAACCGTTCGCTTTCGCTTTCGCTTTCGCTTTC 7693 QY
7625 GCATGGCTCGGAGGAAAGTAAAGTCCCATCAACCGTTCGCTTTCGCTTTCGCTTTCGCTTTC 7684 Db
7694 TCACCAACATGCTGTACGCCCAACATCCGCGAGCGCAAGCTCCGCGAGAAAGAGT 7753 QY
7685 CCACCATAAATGGTTTATGCCACACATCTCGCAGCGCAGGCTTCGCGAGAAAGAGT 7744 Db
7754 CACTTTGACAGATTCGATCTTGGATGATCAATTACCGGACGCTATCAAGGAGATGAA 7813 QY
7745 CACTTTGACAGATTCGATCTTGGATGATCAATTACCGGACGCTATCAAGGAGATGAA 7804 Db
7814 GCGAGGCGTCCACAGTTAAGCTTAAAGTCTTATCTATAGAGGAGCGCTGCAAGCTGAC 7873 QY
7805 GCGAGGCGTCCACAGTTAAGCTTAAAGTCTTATCTATAGAGGAGCGCTGCAAGCTGAC 7864 Db
7874 GCGCGCATTTTCGCGCAAAATCCAAATTTGGCTATGGGCAAAAGGACGCTCCGGAACCTATC 7933 QY
7865 GCGCGCATTTTCGCGCAAAATCCAAATTTGGCTATGGGCAAAAGGACGCTCCGGAACCTATC 7924 Db
7934 CAGCAGGCGCTTAAACACATCCGCTCGTGTGGAGGACTTGTGGAAGACACTGTAAC 7993 QY
7925 CAGCAGGCGCTTAAACACATCCGCTCGTGTGGAGGACTTGTGGAAGACACTGTAAC 7984 Db
7994 ACCAATTTGACACCATCATGGCAAAATGAGGTTTTCTGTGTCCAAACAGAGAGG 8053 QY
7985 ACCAATTTGACACCATCATGGCAAAATGAGGTTTTCTGTGTCCAAACAGAGAGG 8044 Db
8054 AGGCGGAGCGCTCGCTTATCGTATTCAGAGCTGGGAGTTCGCTGATGCGGAA 8113 QY
8045 AGGCGGAGCGCTCGCTTATCGTATTCAGAGCTGGGAGTTCGCTGATGCGGAA 8104 Db

QY 8114 GATGCCCCCTTTACGAGCTGGTCTCCACCCCTTCTCAGCCCGTGATGGGCTCCTCATACGG 8173
DB GATGCCCCCTCTATGATGTTGGTCTCCACCCCTTCTCAGGTGGTGGCTCCTCATACGG 8164
QY 8174 ATTTCAATACTCCCAAGCAGCGGGTCCAGTTCCTGTTGTAATACCTGGAAATCAAGAA 8233
DB ATTCAGTACTCTCTGGGCGAGCGAGTCCAGTTCCTGTTGTAATACCTGGAAATCAAGAA 8224
QY 8234 ATGCCCTATGGCTTCTCATATGACACCGCTGTTTTCAGCTCAACGGTCACTGAGAGTA 8293
DB AAACCCCATGGGCTTTTTCATATGACACTCGCTGTTTTCAGCTCAACGGTCACTGAGAGTA 8284
QY 8294 CATTCGTGTGAGGAGTCAATTTTACCAATGTTGTGACTTGGCCCCCGAGGCGAGACAGGC 8353
DB CATCGTGTGAGGAGTCAATTTTACCAATGTTGTGACTTGGCCCCCGAGGCGAGACAGGC 8344
QY 8354 CATAAAGTTCAGACAGCGGCTTTTCAATCGGGGGTCCCTGACTAACTCAAAAGGGA 8413
DB CATAAATCGCTTCAGACAGCGGCTTTTATATCGGGGGTCTCTGACTAAATCAAAAGGGA 8404
QY 8414 GAACTGCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGAGCTAGCTGGGTAATAC 8473
DB GAACTGCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGAGCTAGCTGGGTAATAC 8464
QY 8474 CCTCATGTACTTTGAAAGGCCACTGCGAGCTGTGAGCTGCAAGCTCCAGGACTGCAC 8533
DB CCTCATGTACTTTGAAAGGCCCTTGCAGCTGTGAGCTGCGAGCTCCAGGACTGCAC 8524
QY 8534 GATGCTGTGAAAGGAGCAGCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAGGAGA 8593
DB GATGCTGTGAAAGGAGCAGCTTGTGCTTATCTGTGAAAGCGCGGAAACCCAGGAGA 8584
QY 8594 TCGCGGGCCCTTACGAGCTTCAAGGAGCTATGACTAGTATTCGCGCCCCCGGGGA 8653
DB CGCGGAGCTTACGAGCTTCAAGGAGCTATGACTAGTATTCGCGCCCCCGGGGA 8644
QY 8654 TCGCGCCCAACAGAAATACGAGCTGAGCTGATAACATCATGTTCTCAATGTGTCAGT 8713
DB CCCGCCCAACAGAAATACGAGCTGAGCTGATAACATCATGTTCTCAATGTGTCAGT 8704
QY 8714 CGCGCAGATGATCTGGCAAAAGGATATACCTTACCTCACCGGTGACCCCAACCCCTCT 8773
DB CGCGCAGATGATCAGGCAAAAGGATATACCTTACCTCACCGGTGACCCCAACCCCTCT 8764
QY 8774 TGCAGGGCTGCTGGGAGACAGCTAGACACTCAATCAACTCTTGGCTAGGCAATAT 8833
DB AGCAGGGCTGCTGGGAGACAGCTAGACACTCAATCAACTCTTGGCTAGGCAATAT 8824
QY 8834 CATCATGTATGGCCCAACCTTATGGCAAGGATATCTGTGATGACTCACTTTTCTCCAT 8893
DB TATCATGTATGGCCCAACCTTATGGCAAGGATATCTGTGATGACTCACTTTTCTCCAT 8884
QY 8894 CTTCTAGCTCAAGAGCAACTTTGAAAGGCTTGGATGTCAGATCTACGGGGCTTGCTA 8953
DB CTTCTAGCTCAAGAGCAACTTTGAAAGGCTTGGATGTCAGATCTACGGGGCTTGCTA 8944
QY 8954 CTCATTGAGCCACTTACCTACCTCAGATCAATGAAAGCTCCATGCTTTAGGCGATT 9013
DB CTCATTGAGCCACTTACCTACCTCAGATCAATGAAAGCTCCATGCTTTAGGCGATT 9004
QY 9014 TACACTCCACAGTACTCTCCAGGTGAGATCAATAGGGTGGTTCATGCTCAGGAACT 9073
DB TACACTCCACAGTACTCTCCAGGTGAGATCAATAGGGTGGTTCATGCTCAGGAACT 9064
QY 9074 TGGGGTACCACTTTCGCAAGCTTGGAGACATCGGGCAGAGTGTCCGCGCTTAAGTACT 9133
DB TGGGGTACCACTTTCGCAAGCTTGGAGACATCGGGCAGAGTGTCCGCGCTTAAGTACT 9124
QY 9134 GTCCAGGGGGGGGGGGCCCACTTGTGGAGATACCTTTTAACTGGGAGTAAGAC 9193
DB GTCCAGGGGGGGGGGGCCCACTTGTGGAGATACCTTTTAACTGGGAGTAAGAC 9184
QY 9194 CAAGCTTAACTCACTTCCCAATCCCGCCGCGTCCCAGCTGGAGTGTCTGGCTGGTTCGT 9253

DB 9185 CAAACTTAACTCACTCAATCCCGGCTGCGTCCCGGCTGAGACTTGTCCGGCTGTTGCT 9244
QY 9254 CGCTGGTTACAGCGGGGAGACATATACAGACTGTCTCGTCCCGACCCCGCTGGTT 9313
DB TGCTGGTTACAGCGGGGAGACATATACAGACTGTCTCGTCCCGACCCCGCTGGTT 9304
QY 9314 TCCGTGTGCTACTCTCTACTCTTCTGTAGGGGTAGGCAATTTACCTGCTCCCAACCGATG 9373
DB CATCTGTGCTACTCTCTACTCTTCTGTAGGGGTAGGCAATTTACCTGCTCCCAACCGATG 9364
QY 9374 AACGGGAGCTAACACTCCAGGCTTAAAGCAATTTCTGTTTTTTTTTTTTT 9425
DB AACGGGAGATAAACACTCCAGGCAATAGGCCATCCCCCTTTTTTTTTTTT 9416

RESULT 8
US-08-150-204E-96
; Sequence 96, Application US/08150204E
; Patent No. 6538126
; GENERAL INFORMATION:
; APPLICANT: CHO, Joong Myung
; LEE, Yong Beom
; PARK, Young Woo
; LIM, Kook Jin
; CHOI, Deog Young
; SO, Hong Seob
; KIM, Chun Hyung
; KIM, Sung taek
; YANG, Jae Young
; TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YANG, Jae Young
; STREET: 386-1, Doryong-dong, Yuseong-gu
; CITY: Daejeon
; STATE: Daejeon
; COUNTRY: Republic of Korea
; ZIP: 305-340
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette, 3.5inch, 1.44MB storage
; COMPUTER: IBM PC/pentium
; OPERATING SYSTEM: Windows
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/150,204E
; FILING DATE: 20-Apr-1994
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: KR 91-9510
; FILING DATE: 10-JUN-1991
; APPLICATION NUMBER: KR 91-13601
; FILING DATE: 6-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Shahar Islam, Esq.
; REGISTRATION NUMBER: 32,507
; REFERENCE/DOCKET NUMBER: 2695/FLK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 940-8564
; TELEFAX: (212) 940-8776
; INFORMATION FOR SEQ ID NO: 96
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; OTHER INFORMATION: KHCV-LBC1, Fig. 2
; SEQUENCE DESCRIPTION: SEQ ID NO: 96
US-08-150-204E-96

Query Match 84.4%; Score 8098.2; DB 3; Length 9472;

Best Local Similarity 90.9%; Pred. No. 0; Mismatches 858; Indels 0; Gaps 0;			
Matches 8613; Conservative 0;			
QY	1	GCAGCCCCCTCATGGGGGCGACACTCCACCATGAATCACTCCCTCTGTGAGGAACCTACTG	60
Db	2	GCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTCTGTGAGGAACCTACTG	61
QY	61	TTTTCACGAGAAAGCGTTACGCATGGCGTTAGTATGATGTGCTGAGCCTCCAGGAC	120
Db	62	TTTTCACGAGAAAGCGTTACGCATGGCGTTAGTATGATGTGCTGAGCCTCCAGGAC	121
QY	121	CCCCCTCCGGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG	180
Db	122	CCCCCTCCGGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG	181
QY	181	GACGACCGGGTCTCTTTCTTGGATCAACCGCTCAATGCCCTGGAGATTTGGGGGTGCCCCC	240
Db	182	GACGACCGGGTCTCTTTCTTGGATCAACCGCTCAATGCCCTGGAGATTTGGGGGTGCCCCC	241
QY	241	GGGAGACTGTAGCCGAGTAGTGTGGGTGCGAAAGGCTTTGTGTACTGTGCTGATAGG	300
Db	242	GGGAGACTGTAGCCGAGTAGTGTGGGTGCGAAAGGCTTTGTGTACTGTGCTGATAGG	301
QY	301	GTGCTTGCAGATGCCCGGAGGTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC	360
Db	302	GTGCTTGCAGATGCCCGGAGGTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC	361
QY	361	CTCAAGAAAAACCAACAGTAAACCAACCGCGCCGACAGGACGTCAAGTTCGCGGGCG	420
Db	362	CTCAAGAAAAACCAACAGTAAACCAACCGCGCCGACAGGATTAAGTTCGCGGGCG	421
QY	421	GTGCTCAGATCGTTGTGTGGAGTTTACCTGTTCGCGCGCAGGGGCCCAAGTTGGGTGTC	480
Db	422	GTGCTCAGATCGTTGTGTGGAGTTTACTTTGTTCGCGCGCAGGGGCCCAAGTTGGGTGTC	481
QY	481	GCAGACTAGGAAGCTTCGAGGGTTCGAACTGTGTGAAAGGCGACCACTATCCCAA	540
Db	482	GCAGACTAGGAAGCTTCGAGGGTTCGAACTGTGTGAAAGGCGACCACTATCCCAA	541
QY	541	AGCTCGCCGACCGAGGCGAGGCTGGGTGAGCCCGGATACCTTTGGCCCTCTATG	600
Db	542	AGCTCGCCGACCGAGGCGAGGCTGGGTGAGCCCGGATACCTTTGGCCCTCTATG	601
QY	601	GCAATAGGCGCTGGGGTGGGAGGATGGTCTCTGTCAACCCGCGGTCCCGGCTAGTT	660
Db	602	GCAATAGGCGCTGGGGTGGGAGGATGGTCTCTGTCAACCCGCGGTCCCGGCTAGTT	661
QY	661	GGGCCCCCAGGACCCCGCGGTAGGTGCGGTAACTTTGGGTAGGTCAATGATACCTTA	720
Db	662	GGGCCCCCAGGACCCCGCGGTAACTTTGGGTAGGTCAATGATACCTCA	721
QY	721	CATCGGCTTCGCGATCTCATGGGGTACATTCGCTCGCGGCCCGCCCTAGGGGGCG	780
Db	722	CATCGGCTTCGCGATCTCATGGGGTACATTCGCTCGCGGCCCGCCCTAGGGGGCG	781
QY	781	CTGCGAGGGCTTCGACACCGGTTCGCGGTTCGAGGACGCGGTGAATATCAACACAG	840
Db	782	TTGCGAGGGCTTCGACACCGGTTCGAGGACGCGGTGAATATGCAACAG	841
QY	841	GGAACTTTCGCGGTTCCTTTCTCTATTCCTCTTGGCTCTGCTGTCTGTGTTGACCA	900
Db	842	GGAACTTTCGCGGTTCCTTTCTCTATTCCTCTTGGCTCTGCTGTCTGTGTTGACCA	901
QY	901	TCCAGCTTCGCTTATGAATGGCGAAGTGTCCGGATATACATGTACGACGAGCT	960
Db	902	CCCCAGTTTCGCTTATGAATGGGTAAACGGGTCCGGATATACATGTACGACGAGCT	961
QY	961	GCTCCAACTCAAGCATTTGTATGAGGAGGAGGTGATGATGATCTCCCGGGTGC	1020
Db	962	GCTCCAACTCAAGCATTTGTATGAGGAGGAGGTGATGATGATCTCCCGGGTGC	1021
QY	1021	TGCCCTGTGTTCAGGAGGTTAAAGCTCCCGTGTGTGGGTAGCGCTCACTCCCAAGCTCG	1080
Db	1022	TGCCCTGTGTTCAGGAGGTTAAAGCTCCCGTGTGTGGGTAGCGCTCACTCCCAAGCTCG	1081
QY	1081	CGGCCAGGAATGCCAGCGTCCCACTAGCAATAGAGCGCCACCTCGACTTGTCTGTTG	1140
Db	1082	CGGCCAGGAATGCCAGCGTCCCACTAGCAATAGAGCGCCACCTCGACTTGTCTGTTG	1141
QY	1141	GGAGCGCTGTCTTCTGCTCCGCTATGTATGTAGTGGGGATCTCTGCGGATCTATTTCTCG	1200
Db	1142	GGGTAGCTGTCTTCTGCTCCGCTATGTAGTGGGGATCTCTGCGGATCTATTTCTCTG	1201
QY	1201	TCTCCAGCTGTTCACCTTCTCGCTCGCGGATGAGACAGTGCAGGACTGCAACTGCT	1260
Db	1202	TCTCCAGCTGTTCACCTTCTCGCTCGCGGATGAGACAGTGCAGGACTGCAACTGCT	1261
QY	1261	CAATCTATCCCGGATGTATCAGGTCAACCGCATGCTTGGGATATGATGATGAACTGCT	1320
Db	1262	CAATCTATCCCGGCGCGTATCAGGTCAACCGCATGCTTGGGATATGATGATGAACTGCT	1321
QY	1321	CACCTAACACAGCCCTAGTGTGTGCGAGTGTCTCGGATCCCAAGCTGTCTGTGACA	1380
Db	1322	CGCTTAAACAGCCCTAGTGTGTGCGAGTGTCTCGGATCCCAAGCTGTCTGTGACA	1381
QY	1381	TGCTGCGCGGGGCCCACTGCGGAGTCTCTGCGGCTTGTGCTACTATTCATGTTAGGA	1440
Db	1382	TGCTGACAGGCTCCCACTGCGGAATCTCTGCGGCTTGTGCTACTATTCATGTTAGGA	1441
QY	1441	ACTGGCTTAAGGTTCTGATTTGGCGTACTCTTTGCGGCGTGTGACGGAGACCCACA	1500
Db	1442	ACTGGCTTAAGGTTCTTAAATTGCGATGCTACTCTTTGCGGCGTGTGACGGAGACCCACG	1501
QY	1501	CGAGGGGAGGGTGGCGGCCACACACCTCGGGTTCAGTCCCTTTCTCATCTGGGG	1560
Db	1502	TGACAGGGGGGGCGAAGGTCGGGCGCTAGCTGCTGCTAAGTCCCTTTTAGCCCTGGGC	1561
QY	1561	CGTCTCAGAAAAATCCAGCTTGTGAATACCAACCGAGCTGGCACATCAACAGAGTGC	1620
Db	1562	CGTCTCAGCACCTCCAGCTCATAAACCAACCGAGCTGGCATATCAACAGGAGCCGCC	1621
QY	1621	TAAATGCAATGATCTCCCTCAAACTGGGTCTTTGCGGCGCTGTTTACGCACACAAGT	1680
Db	1622	TGAGCTGCAATGATCTCCCTCAACACTGGGTCTTTGCGGCGCTGTTTACAAAATACAGT	1681
QY	1681	TCAACTGTTCGGGTGGCGGCGGATGCGGAGCTGCGGCCCATTCAGTCTGTTGCGCC	1740
Db	1682	TCAACGCTCGGGTGGCGGAGCGCTTGGCCACAGTGCGCCCCCATTCATACATTCGCGC	1741
QY	1741	AGGGTGGGGCCCATCACCTATATAAGCTTAAAGCTTAAAGCTCGGATCAGAGCCTTATTGCT	1800
Db	1742	AGGGTGGGGTCCCATCACCTTACCTAGGCTCATGATTTGGATCAGAGGCCCTATTGCT	1801
QY	1801	GGCATTAACGCGCTTCGACCGCTGTGGTGTCTGTAACCGGCTCGAGGTGTGGTCCAGTGT	1860
Db	1802	GGCATTAACGCGCTTCAACCGCTGTGGTGTCTGTAACCGGCTCGAGGTGTGGTCCAGTAT	1861
QY	1861	ATTGTTTACCCCAAGCCCTGTGTGGTGGGGACCGGATCGTTCCGCTGCTCCTAGT	1920
Db	1862	ACTGCTTACCCCGAGTCTCTTTCGGTGGGGATACCGATCGTTCGGTCCCTCAT	1921
QY	1921	ATAGCTGGGGGAGATGAGACAGCTGTGCTCTCAACCAACAGCGCTCCGCGCAACAG	1980
Db	1922	ACAGTGGGGGGAATGAGACAGGAGCTGTGCTCTTAAACAGCGCGGGCGCGGCAAG	1981
QY	1981	GCAACTGTTTCGGCTGTACATGATGAATAGTACTGGGTTCATTAAGACGTGCGGAGTGC	2040
Db	1982	GCAACTGTTTCGGCTGTACATGATGAATAGTACTGGGTTCATTAAGACGTGCGGAGTGC	2041
QY	2041	CCCCGTGTAAACATCGGGGGGTTCGGTAAACCGCACTTGTGATTCGCCCAAGCACTTCC	2100
Db	2042	CCCCGTGTAAACATCGGGGGGTTCGGTAAACCAATACCTTTGACCTGCCCAAGCACTTCC	2101
QY	2101	GGAGACCCCGAGGCTTACACAAATGTGCTGGGGCCCTGGTTGACACCTAGT	2160
Db	2102	GAAAGCACCCCGGGGCCATTTACACCAATCGGGTTCGGGGCCCTGGTTTAAACACCGAGT	2161

Qy	2161	GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCT	2222
Db	2162	GCTTAGTCGACTACCCGCTACAGGCTCTGGCATTTACCCCTGCACTGTCAACTTTTACCATCT	2221
Qy	2221	TTAAGGTTAGGATGTATGTGGGGGGCGTGAGACAGAGCTCAATGCGGCATGCAATTGGA	2280
Db	2222	TTAAGGTTAGGATGTATCGTGGGGGGCGGAGACAGAGCTCGAGCGCGATGCAACTGGA	2281
Qy	2281	CTCGAGAGAGCGCTGTAACTTCGGAGACAGGAGTAGGTGAGAACTCAGCCGCTGCTGC	2340
Db	2282	CTCGGAGAGCGTGTGTGA	2341
Qy	2341	TGCTCAACAGAGTGGCAGATACGCGCTGTGTCTTTCACACCTACCGCTTTATCCA	2400
Db	2342	TGTCTCAAAGAGTGGCAGGTACTGCCCTGTCTTTCACAAACCTACCGCTCTGTCCA	2401
Qy	2401	CTGGTTTGATCCATCTCCATCAGAAACATGTAAGCGTGAACATACCTGTACGGTGTAGGGT	2460
Db	2402	CTGGTTTGATTCATCTCCATCAGAAACATGTAAGCATACAAATACCTGTACGGTATAGGGT	2461
Qy	2461	CAGCGTTGTCTCCTTTGGCAATCAAATGGAGGTACATCCTGTGTCTTTTCCTTCTCTGG	2520
Db	2462	CGGCGGTTGTCTCCTTTGGCATCAAATGGAGGTATATTGTGTCTCTTCTCTTCTCTGG	2521
Qy	2521	CAGACGCGCGTGTGTGCTGTGTGTGGATGATGCTGTGATAGCCAGCGCTGAGGCGG	2580
Db	2522	CGACGCGCGGCTCTGCGCTGTGTGTGTGGATGATGCTGTGTGTAGCGACGCGAGGCGG	2581
Qy	2581	CCTTAGAGAACTTGGTGGTCTCTCAATCGCGCGTCTGTTGGCGGAGCGCATGTATTCTCT	2640
Db	2582	CCTTAGAGAACTCGTGGTCTCTCAATCGAGCGTCCGTGGCGGAGCGCATGTCAATCTTT	2641
Qy	2641	CCTTTCTTTGTGTTCTTCTGCGCGCGCTGTGTACATTAAGGCGAGGCTGGCTCTGGGGCGG	2700
Db	2642	CCTTCAATTGTGTTCTTCTGTGCTGCTCTGTGTACATCAAGGCGAGGCTGGTTCGCGAGCGG	2701
Qy	2701	CGTATGCTTTTATGGCGTATGGCGCGTGTCTCTGCTCTTACTGGCGGTACCAACACGAG	2760
Db	2702	CATACGCCCTCTATGGCGTATGGCGCGTCTGTCTGTCTGTGGCGTATACCAACGCGG	2761
Qy	2761	CTTAGCGCTTGGACCGGGAGATGGCTGCATCGTGGGGGGTGGGTTCTTTGTAGTCTGG	2820
Db	2762	CGTAGCCATGACCGGGAGATGACCGCATCGTGGGAGGCGGTTTTTGTAGTCTGG	2821
Qy	2821	TATTCTTGAACCTTGTCACCATACATAAAGTGTTCATCTAGGCTCATATGGTGTTTAC	2880
Db	2822	TACTCTTGACCTTGTCAACCACTATAAAGTGTTCCTTGGCAGGTTTCATATGGTGGCTAC	2881
Qy	2881	AATACTTTATCAGAGAGCGGAGCGACATGCAAGTGTGGTCCCGCCCTCAACGTTTC	2940
Db	2882	AATATCTCATCACCAGAACCGAGCGCATCTGCAAGTGTGGTCCCCCTCTCAACGTTTC	2941
Qy	2941	GGGAGGCGCGATGCCATCATCTCTCTCAGCTGTGGGTTTCATCCAGAGTTAATTTTG	3000
Db	2942	GGGGGGTTCGGAATGCCATCATCTCTCATATCGTGGTCCACCCAGAGCTAATCTTTG	3001
Qy	3001	ACATCAACAAACTCCTGCTGCCATATCTGGCGCGCTCATGTGTGCTCCAGGCTGGCATAA	3060
Db	3002	ACATCAACAAATAATTGCTGCCATATTGGCGCGCTCATGTGTGCTCCAGGCGGCATAA	3061
Qy	3061	CGAGATGCCGTACTTCTGTGGCGGCTCAAGGGCTCATTCGTGCATGATGTTAGTGGCAA	3120
Db	3062	CTAGATGCCGTACTTCTGTGGCGGCAAGAGGGCTCATTCGTGTGATGATGTTGGCGCGGA	3121
Qy	3121	AAGTCGCGGGGGTCAATTATGTCCAAATGGTCTTCATGAAGCTGGGCGCGCTGACAGGTA	3180
Db	3122	AAGTCGTGGGGGTCAATTACGTCCAAATGGTCTTCATGAAGCTGGCGGCACTAGCAGGTA	3181
Qy	3181	CGTAGGTTTATAACCATCTTATCCCACTGGGGGACTGGGCCACCGCGGGCTACGAGACC	3240
Db	3182	CGTAGGTTTATGACCATCTTACTCCACTGCGAGATTGGGCTCAACACGGGCTTACGAGACC	3241

Qy		3241	TTCCGGTGGCGGTAGAGACCGCGTCGTCCTTCTCCGCCATAGAGACAAGGTCAATCACCTGGG	3300
Db		3242	TTGCAGTGGCAGTAGAGACCGGTGTCTTCTGACATGGAGACCAAAGTCATCACCTGGG	3301
Qy		3301	GAGCAGACACCGCTGCGGTGGGGACATCATCTTGGGTCTTACCCTCTCCGCCCGGAAGG	3360
Db		3302	GGGCAGACACCGCGCGGTGGGGGACATCATCTTGGCTGCCCTGCTTTCGCCCGGAAGG	3361
Qy		3361	GGGAAGGAGATATTTTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTGGCGACTCCTTG	3420
Db		3362	GGGAAGGAGATATTCTGGGACCGGCCGATAGTCTTGAAGGACAGGGGTGSCGACTCCTTG	3421
Qy		3421	CGCCCATCAGCGCTACTCTCCCAAACAAACGCGAGCGCTGCTTGGTTGTCATCATCACTAGCC	3480
Db		3422	CGCCCATCAGCGCTACTCTCCCAAACAAACGCGAGCGCTGCTTGGTTGTCATCATCACTAGCC	3481
Qy		3481	TCCAGCGCGGACAGAACACAGGTGGAAGGGAGGTTCAAAGTGGTTCTTACCAGCAACAC	3540
Db		3482	TTACAGCGCGGACAGAACACAGGTGAGGGGAGGTTCAAAGTGGTTCTTACCAGCAACAC	3541
Qy		3541	AATCTTCTTCGGCGACCTGTCATCAAACGGCGTGTCTGGACTGTCTACCATGGCGCTGGCT	3600
Db		3542	AATCTTCTTCGGCGACCTGTCATCAAAGCGTGTGTGGACTGTCTTCCACGGCGCGGCT	3601
Qy		3601	CGAAGACCTTAGCGGTCFAAAGGTCCAATACACCAAAATGTATACAAATGTATAGACTGG	3660
Db		3602	CMAAGACCTTAGCGGCCAAAGGTCNAATCACCCAAATGTATACAAATGTATAGACCAGG	3661
Qy		3661	ACCTCGTGGCTGGCAGGCGCCCCCGGGGCGCTCCATGACACCATGCAGCTGTGGCA	3720
Db		3662	ACCTTGTGTGTGGCGGGACCTCTCTGGGGCGGTTCCCTGACACCATGCATTTGCGGCT	3721
Qy		3721	GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCAATTCGGGTGCGCGCGGAGCGG	3780
Db		3722	CCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCAATTCGGGTGCGCGGGGGGTG	3781
Qy		3781	ACAGCGGGGAAGTCTACTCTCCCCAGGCGCGTCTCTACTCTCTGAAAAGGTCTCTCGGGT	3840
Db		3782	ACGGTAGGGGAGCGCTACTCCCCCGGAGCGCTGTCTCTACTTGAAGGGCTCTCTCGGGT	3841
Qy		3841	GTCCATTTGCTTTGCCCTTGGGGGACGTCTGTGGCGCTCTTCCGGGCTGCTGTGCAACC	3900
Db		3842	GTCCACTGTCTGCCCCTTGGGGGACGCTGTGGGCACTCTTCCGGCTGTGTATGCAACC	3901
Qy		3901	GGGGGTGCGGAAGCGGTGGACTTCATACCCGTGTAGTCTATGGAAATACCATGCGGT	3960
Db		3902	GGGGGGTTGCCATGGCGGTGGAAATTCATACCCGTGTAGTCTATGGAAATCTATGCGGT	3961
Qy		3961	CTCCGGTCTTACAGACAACTCAACCCCCCGGCTGTACCGCAGACATTCOAAGTGGCAC	4020
Db		3962	CTCCGGTCTTACAGGACAAATCCGTCTCCCCCGGCTGTACCGCAGACATTCOAAGTGGCCC	4021
Qy		4021	ATCTGACGCTCTACTTGGCAGCGGCAAGACCAAAAGTGGCGGCTGCGTATGCAAGCC	4080
Db		4022	ACTTACACGCTCCCAACCGGACGGGCAAGACACTAGGGGTGCGGGCTGTCATATGCAGCCC	4081
Qy		4081	AAGGGTCAAGGTGCTCGTCTGAACCCGTCCGTTGCCGCCACCTTAGGGTTTGGGGCGT	4140
Db		4082	AAGGGTCAAGGTGCTCGTCTTAATCCGTTCCGTCGCGCGGCACTTCTGGGTTTTGGGGGT	4141
Qy		4141	ATATGTCCAAAGGACACAGGTATCGACCTTAACATCAGAACTGGGGTGAAGCACTTACCA	4200
Db		4142	ATATGTCCAAAGGACATGGTATCGACCCCAACCTTAGAACTGGGGTTAAGGACCATCACCA	4201
Qy		4201	CGGGGGCTTCATTAGCTATCTCAACCTTATGGCAAGTTCTTGGCGACGGTGGCTGTCTG	4260
Db		4202	CAGGTGCCCTTATCACATCTTCCACCTTATGGCAAGTTCTTGGCGACGGTGGGGGCTCCG	4261
Qy		4261	GGGGCGCTATGACATCATATAATGTGATGAGTGCCACTCAACTGACTCGACTACCATCT	4320
Db		4262	GGGGCGCTATGACATCATATAATGTGATGAGTGCCACTCAACTGACTCGACTACCATCTT	4321
Qy		4321	TGGGCATCGGCACAGTCTCTGBAACCAAGCGGACGCGTGGAGCGCGCTGTGTGTGTCTG	4380

Db		4322	ATGGCATCGSCACAGTCTCTGSCCAAGCGGAGACGGCTGGAGCGGGCTCGTGGTGTCTCT		4381
Qy		4381	CCACCGCTACACTCCGGGATCGGTACCGTGCACACCCCAATATCGAGGAAATAGGCC		4440
Db		4382	CCACCGCTACCGCTCCGGGATCGGTACCGTGCACACCTCAATATCGAGGAGTGGCC		4441
Qy		4441	TGTCCAAATAGAGAGATCCCTTCTATGGGAAAGCCATCCCATTTAGAGGCCATCAAG		4500
Db		4442	TGTCTAATACTGGAGAGATCCCTTCTACGGCAAGCCATCCCATTCGAGGCTATCAAG		4501
Qy		4501	GGGGAGGATCTCATTTCTGCATTCGAAGAAATGTGACGAGCTCGCCGCAAGC		4560
Db		4502	GGGGAAGGCATCTCATTTCTGCCATTCGAAGAAAGTGTGACGAACTCGCCGCAAGC		4561
Qy		4561	TGACAGCCTCGAGCTGAACGCTGTAGCATATTAACCGGGGCCCTTGATGTCCGTCATAC		4620
Db		4562	TGTCAGCCTCGAGCTCAATGCGGTAGCGTATTAACGGGGCTTGACGTGTCGTCATAC		4621
Qy		4621	CGCCTATCGGAGACGTGTTGTGCTGGCAACAGACGCTCTAATGACGGGTTTACCGGCG		4680
Db		4622	CGACAGCGGAGACGTGTTGTGCTGGCGACGAGCGCTCTAATGACGGGCTTTACCGGCG		4681
Qy		4681	ATTTGACTGAGTGCAGTCAATACATGTTGTACCCAGACAGTCGACTTCAGCTTGG		4740
Db		4682	ACTTTGACTGAGTGCAGTGTAAATACGTGTGTCAACCCAGACAGTCGATTTGAGCTTGG		4741
Qy		4741	ATCCCACTTACATTTGACGACACACCGTGTGCCAAGACGGGTGTCCGCTCGCAAC		4800
Db		4742	ACCCCACTTACATTTGACGACACACCGTGTGCCAAGACGGGTGTCCGCTCGCAGA		4801
Qy		4801	GGCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC		4860
Db		4802	GGCGAGGACGACTGTTAGGGGAGGGCTGGCATATACAGGTTTGTGACTCCAGGAGAAC		4861
Qy		4861	GGCCCTCGGGCATGTCGATCTTCGCTCTGTGTGAGTCTATGACGGGGCTGTGCTT		4920
Db		4862	GGCCCTCGGGCATGTCGATCTTCGCTCTGTGTGAGTCTATGACGGGGGTGTGCGT		4921
Qy		4921	GGTATGAGCTCAGCCCGCTGAGACCTCGGTAGTTAGTTGGGCTTACTTAAATACACAG		4980
Db		4922	GGTACGAACTCAGCCCGCTGAGACCTCGGTAGTTAGTTGGGCGGTACTTAAACACACAG		4981
Qy		4981	GGTTGCCGCTCTGCCAGGACCATCTGAGTTCTCGGAGGGTGTCTTACAGGCGCTACCC		5040
Db		4982	GGTTGCCGCTCTGCCAGGACCATCTGAGTTCTCGGAGGGTGTCTTACAGGCGCTACCC		5041
Qy		5041	ACATAGATGCCCACTTCTTATCCAGACTAAACAGGCGAGAGAACTTCCCTTACTTGG		5100
Db		5042	ACATAGATGCCCACTTCTTATCCAGACTAAACAGGCGAGAGAACTTCCCTTACTTGG		5101
Qy		5101	TGGCATATCAAGCTACAGTGTGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAA		5160
Db		5102	TAGCATATCAAGGCTACAGTGTGCCAGGGCTCAAGGGCTCAAGGGCTCAAGGGCTCAAG		5161
Qy		5161	TGTGGAAGTGTCTCATACGGCTGAAAACCTACACTGCAAGGGGCAACACCCCTGCTGATA		5220
Db		5162	TGTGGAAGTGTCTCATACGGCTGAAAACCTACACTGCAAGGGGCAACACCCCTGCTGATA		5221
Qy		5221	GGCTAGGAGCGGTCCTCAAAATGAGGTCTATCTTCAACACCCCATAACTTAAATCATG		5280
Db		5222	GGTTAGGAGCGGTCCTCAAAATGAGGTCTATCTTCAACACCCCATAACTTAAATCATG		5281
Qy		5281	CATGATGTGGGTCGACCTGAGGTGCTGCTAGGACCTGGGTGCTGAGGCGGAGTCC		5340
Db		5282	CATGATGTGGGTCGACCTGAGGTGCTGCTAGGACCTGGGTGCTGAGGCGGAGTCC		5341
Qy		5341	TTGAGGCTTTGGCGCATACCTGCTGACGACGAGTGTGCTGCTGCTGCTGCTGCTGCTG		5400
Db		5342	TCCAGGCTCTGGCGCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG		5401
Qy		5401	TCTTGTTCGGGAAGCCAGCTGCTGCTTCCGACAGGGAAGTCTCTTACAGGAGTTCGATG		5460

Db		5402	TCCTGTCCGGAAGCCGGCTATCATCCCGATAGGGAAGTTCTTACCAGGAGTTCCAGC		5461
Qy		5461	AGATGGAAGAGTGTGCTCTCAAACTTCTTATCATCGAGCAGGAATGCAAGCTCGCCGAGC		5520
Db		5462	AGATGGAAGAGTGTGCTCTCAAACTTCTTATCATCGAGCAGGAATGCAAGCTCGCCGAGC		5521
Qy		5521	AATTCAAGCAAAAGCGCTCGGGTGTGTCAAAACGGCCACCAAGCAAGCGGAGGCTGCTG		5580
Db		5522	AATTCAAGCAAAAGCGCTCGGGTGTGTCAAAACGGCCACCAAGCAAGCGGAGGCTGCTG		5581
Qy		5581	CTCCGCTGTGAGTCCAAAGTGGGAGCCCTTGAAGACCTTCTGGGCGAAGCAATGTGGA		5640
Db		5582	CTCCGCTGTGAGTCCAAAGTGGGAGCCCTTGAAGACCTTCTGGGCGAAGCAATGTGGA		5641
Qy		5641	AATTCAAGCAAAAGCGCTCGGGTGTGTCAAAACGGCCACCAAGCAAGCGGAGGCTGCTG		5700
Db		5642	ACTTCAATAGTGGGATACAGTACTTGGGAGGCTTGTCCACTTGTCCCTGGGAACCCCGCAA		5701
Qy		5701	TAGCATATTGATGGCATTTACAGCTTCTATCACTAGCCCGCTCACACCCCAAAACACCC		5760
Db		5702	TAGCATATTGATGGCATTTACAGCTTCTATCACTAGCCCGCTCACACCCCAAAACACCC		5761
Qy		5761	TCTGTGTTTAAACATCTTGGGGGAGTGGTGGCTGSCCAACTCGCTCTTCCAGCGCTGCTG		5820
Db		5762	TCTGTGTTTAAACATCTTGGGGGAGTGGTGGCTGSCCAACTCGCTCTTCCAGCGCTGCTG		5821
Qy		5821	CAGCTTTCGTGGGCGCGGATCGCCGAGCGGCTGTGGCAGCATAGGCTTTGGGAAGG		5880
Db		5822	CAGCTTTCGTGGGCGCGGATCGCTGGAGCGGCTGTGGCAGCATAGGCTTTGGGAAGG		5881
Qy		5881	TGCTCTGGGACATCTTGGGGGCTATGGGGAGGGGTAGCCGGCGCACTCGTGGCCCTTTA		5940
Db		5882	TGCTCTGGGACATCTTGGGGGCTATGGGGAGGGGTAGCCGGCGCACTCGTGGCCCTTTA		5941
Qy		5941	AGTCTATGAGCGGAGGTCCTCCACCGAGGACCTGTGTCAACTTACTCCCTGCCATCC		6000
Db		5942	AGTCTATGAGCGGAGGTCCTCCACCGAGGACCTGTGTCAACTTACTCCCTGCCATCC		6001
Qy		6001	TCTCTCTGTGTGCTCGTGGGGTGTGTGGGAGCAATACTCGCTCGGCGAGTGG		6060
Db		6002	TCTCTCTGTGTGCTCGTGGGGTGTGTGGGAGCAATACTCGCTCGGCGAGTGG		6061
Qy		6061	GCCGGGAGGAGGGGCTGTGAGTGAACCGGCTGTAGGCTTCCGCTTCCGCGGGTA		6120
Db		6062	GCCGGGAGGAGGGGCTGTGAGTGAACCGGCTGTAGGCTTCCGCTTCCGCGGGTA		6121
Qy		6121	ACCAGCTCTCCCTAGCAGCTATGTGCTCAGAGCGAGCGCTGCGAGCGGCTTACCAGA		6180
Db		6122	ACCAGCTCTCCCTAGCAGCTATGTGCTCAGAGCGAGCGCTGCGAGCGGCTTACCAGA		6181
Qy		6181	TCTCTCTAGCCTTACATCACTCACTGCTGAGAGCGAGCTGCGAGCACTGTCTCAGAG		6240
Db		6182	TCTCTCTAGCCTTACATCACTCACTGCTGAGAGCGAGCTGCGAGCACTGTCTCAGAG		6241
Qy		6241	ACTGCTCTAGCCTTACATCACTCGGCTCGGCTGTAGGAGTGTGGGATTTGAGTATGAGG		6300
Db		6242	ACTGCTCTAGCCTTACATCACTCGGCTCGGCTGTAGGAGTGTGGGATTTGAGTATGAGG		6301
Qy		6301	TGTTGACTGACTTCAAGACCTGGCTCCAGTCCAAACTCTTCCCGCGGTTTACCGGAGTCC		6360
Db		6302	TGTTGACTGACTTCAAGACCTGGCTCCAGTCCCAAGTCTTCCCGCGGTTTACCGGAGTCC		6361
Qy		6361	CTTTCTCTCATGCAACGGGGTATCAAGGAGTCTGGCGGGGAGCGGCATCATGACA		6420
Db		6362	CTTTCTCTCATGCAACGGGGTATCAAGGAGTCTGGCGGGGAGCGGCATCATGACA		6421
Qy		6421	CCACTCTCATGCGGAGCAGATCCCGGAGCATGTCAAAAACGGTTCATGAGGATCG		6480
Db		6422	CCACTCTCATGCGGAGCAGATCCCGGAGCATGTCAAAAACGGTTCATGAGGATCG		6481
Qy		6481	TAGGGCTTAGAACCTGAGCAACAGTGGGAGCAAGTTCCTCCCATCAACGATACCA		6540
Db		6482	TGGGCTTAAACCTTGAGCAACAGTGGTACGGGACATTTCCCATCAACGGGTACACA		6541

QY 6541 CGGGACCTTGCACACCCCTCCCGGCGCCAACTATTTCAGGCGCGCTATGCGGGTGGCTG 6600
Db 6542 CGGGCCCTGACACACCCCTCCCGGCGCCAACTATTCAGGGCATTTGGAGAGTGGCGG 6601
QY 6601 CTGAGGAGTACGTGAGAGTTAGCGGTGTGGGGATTTCCACTACGTGACGGGCATGACCA 6660
Db 6602 CTGAGGAGTACGTGAGAGTACGGCGGTGGGAGATTTTCACTACGTGACGGGCATGACCA 6661
QY 6661 CTGACAAAGTAAAGTCCCATGCTCAGAGTTCCGGCCCCCGAATTTCTTACGGAGGTGGATG 6720
Db 6662 CTGACAAAGTAAAGTCCCATGCTCAGAGTTCCGGCCCCCGAATTTCTTACGGAGGTGGATG 6721
QY 6721 GAGTCCGGTTGCACAGGTACGCTCCGGGTGCAAACTCTTCTACGGAGGACGTCAGT 6780
Db 6722 GAGTCCGGTTGCACAGGTACGCTCCGGGTGCGAGACTCTCTTACGGAGGAGGTCTGTAT 6781
QY 6781 TCCAGGTCCGGGCTCAACCAATACTTGGTCCGGTCCAGCTCCCATGCGAGCCGCAACCGG 6840
Db 6782 TCCAGGTCCGGGCTCCACAGTACCTGGTCCGGTCCAGCTCCCATGCGAGCCGCAACCGG 6841
QY 6841 AGTTAAAGTGTACTTCCATGCTCAACGATCCCTCCCATATACAGGAGGACGGCTA 6900
Db 6842 ATGTAGCAGTGTCACTTCCATGCTCACTGACCCCTCCCATATACAGGAGGACGGCTA 6901
QY 6901 AGCGTAGGCTGGCTAGAGGCTCTCCCGCTCTTTAGCCAGCTCATCAGTACGCCAGTTGT 6960
Db 6902 AGCGTAGGCTGGCCAGGGGCTCTCCCGCTCTTTAGCCAGCTCTTTCAGTAGCCAGTTGT 6961
QY 6961 CTGCGCCCTCTTTGAAGGCGCATGCACTATCCACCATGACTCCCGGACGCTGACCTCA 7020
Db 6962 CTGCGCCCTCTTTGAAGGCGCATGCACTATCCCATCATGACTCCCGGACGCTGACCTCA 7021
QY 7021 TCGAGGCCAACTCTTGTGGCGGAGGATGGGCGGAAACATCATCTCGTGGAGTCCAG 7080
Db 7022 TTGAGGCCAACTCTTGTGGCGGAGAGATGGGCGGAAACATCACCCGCGTGGAGTCCAG 7081
QY 7081 AGAATAAGGTAGTAATCTTGGACTCTTTTCGAACCGCTTCACGCGAGGGGATGAGAGG 7140
Db 7082 AGAATAAGGTAGTAATCTTGGACTCTTTTCGACCCGCTCCGAGCGGAGGATGATAGGGGG 7141
QY 7141 AGATATCCGTCGCGCGGAGATCTCGGAAATCCAGGAATTCGCCCTCAGCGTTGCCCA 7200
Db 7142 AAATATCCGTCGCGCGGAGATCTCGGAAATCCAGGAATTCGCCCGAGCGCTGCCCA 7201
QY 7201 TATGGGACGCGCGGACTAACAATCTCTCACTAGAGTCTTGGAGGACCCGGACTACG 7260
Db 7202 TATGGGCGCGCGGATTAACAACCCCTCCGCTGTAGAGTCTTGGAGGACCCGGACTACG 7261
QY 7261 TCCCTCCGTTGTACACGGATGCCATTCGCACTTACCAAGGCTCTCCAAATACCACCTC 7320
Db 7262 TTCTCTCCGTTGTACACGGGTCGCGGTTCGCGCCACCAAGGCCCTCCAAATACCACCTC 7321
QY 7321 CACGAGAAAGAGGACGGTTGTCTTGACAGAATCCAAATGTGTCTTCTGCTTGGCGGAGC 7380
Db 7322 CACGAGGAGAGGACGGTTGTCTTGACAGAATCCACCGTGTCTTCTGCTTGGCGGAGC 7381
QY 7381 TCGCCATAAAGACTTTCGGTAGCTCCGGATCGTGGCCGTTGATACGGGCAACGGGACCG 7440
Db 7382 TCGCTACTAAGACTTTCGGCAGCTCCGGATCGTGGCCATCGACACGGGTAACGGGACCG 7441
QY 7441 CCCTTCTGACTCGGCTCCGAGGAGGTGACAAAGGATCCGACGTTGAGTCTGACTCCT 7500
Db 7442 CCCCTCTGACCAAGCTCCCGGTGACGGGACAGAGAGTCCGACGTTGAGTCTGTTCTCCT 7501
QY 7501 CCATGCCCCCTTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTA 7560
Db 7502 CCATGCCCCCTTTGAGGAGAGCGGGGAGCCCCGATCTCAGCGACGGATCTTGGTCCA 7561
QY 7561 CCGTAGTGAGGAGGCTAGTGAAGATGTCTGTGTGCTCTCAATGTCTTATACGTGGACAG 7620
Db 7562 CCGTAGGAGGAGGCTAGTGAAGAGCTGTCTGTCTGTTTCGATGTCTTACATATGACAG 7621

QY 7621 GCGCCCTGATCACGCCATGCGCTCGGAGGAAAGTAAGTCCCATCAACCCGTTGAGCA 7680
Db 7622 GCGCCCTGATCACGCCATGCGCTCGGAGGAAAGCAAGTTCGCCCATCAACCCGTTGAGCA 7681
QY 7681 ACTCTTGTCTGCGTACCAACAACATGTGTCTAGGCAACAATCCGCGAGCGCAAGCCTCC 7740
Db 7682 ATTCTTGTCTACGTCAACCAACATGTGTCTATGCTACAAATCCGCGACGCGAGCCTGC 7741
QY 7741 GGCAGAAAGAGTCCACTTTTGACAGATTGCAAGTCTCGATGATCATTTACCGGACGTAC 7800
Db 7742 GGCAGAAAGAGTCACTTTTGACAGATGCAAGTCTCGAGCAACAATTTACCGGACGTAC 7801
QY 7801 TCAAGGAGATGAAGCGGAGCGCTCCACAGTTAAGGCTAAAGCTTCTATCTATAGAGGAG 7860
Db 7802 TTAAGGAGATGAAGCGGAGCGCTCCACAGTTAAGGCTAAACTTCTATCTGTAGAAGAAG 7861
QY 7861 CCTGCAAGCTGACGCGCCCAATTTGGGCAAAATCCAAATTTGGCTATGGGCGCAAGGAGC 7920
Db 7862 CTTGCAAACTGACGCGCCCAATTTGGGCAAAATCCAAATTTGGCTACGCGGCGCAAGGAGC 7921
QY 7921 TCCGMACTTATCCAGCAGGCGCTTAACCATCCGCTCGTGTGGAGGACTTGTCTGG 7980
Db 7922 TCCGAGGCTTATCCAGCAGGCGCTTACCCACATCCGCTCGTGTGGAGGACCTTGTCTGG 7981
QY 7981 AAGACACTGAAACCAACCAATTTGACACCACTATGCGCAAAAAGTGAGGTTTTCTCGCTCC 8040
Db 7982 AAGACACTGAAACCAACCAATTTAGCACTACCATCATGGCAAAAATGAGGTTTTCTGTCTCC 8041
QY 8041 AACCAGAGAGGAGGCGCAAGCCAGCTCGCTTATCGTATTTCCAGACTTGGAGTTTC 8100
Db 8042 AACCAGAGAGGAGGCGCAAGCCAGCTCGCTTATCGTGTTCAGATCTGGGAGTTTC 8101
QY 8101 GTGTATGCGAGAAAGTGGCCCTTTACGACGTTCTCCACCTTCTCAGGCGCGTGAAGG 8160
Db 8102 GTGTATGCGAGAAAGTGGCCCTTTATGACGTGGTCTCCACCTTCTCAGGCGCGTGAAGG 8161
QY 8161 GCTCTCTATACGGATTTCAATACTCCCCCAAGCAGCGGTTCGAGTTCTCTGTGTAATACCT 8220
Db 8162 GCTCTCTATACGGATTTCAAGTACTCTCTTAAGCAGCGGTTCGAGTTCTCTGTGTAATACCT 8221
QY 8221 GGAAATCAAAAGAAATGCGCTATGGGCTTCTCATATGACACCGCGTGTGTGACTCAACGG 8280
Db 8222 GGAAATCAAAAGAAATGCGCCATGGCTTCTCATATGACACCGCGTGTGTGACTCAACGG 8281
QY 8281 TCACGTGAGAGTACATTTCTGTGTGAGGAGTCAATTTACCAATTTGACTTGTGGCCCCCG 8340
Db 8282 TCACGTGAGAGTACATTTCTGTGTGAGGAGTCAATTTACCAATTTGACTTGTGGCCCCCG 8341
QY 8341 AGGCCAGACAGGCCATAAGGTCCGCTCACAGACGGCTTTTACATCGGGGGTCCCTGACTA 8400
Db 8342 AAGCCAAACTGGCCATAAAGTCCGCTCACAGACGGCTCTATATCGGGGTCCCTGACTA 8401
QY 8401 ACTCAAAAGGCGAGAACTGGCGTTATCGCGGTGCGCGCAAGTGGCGTGTGACGACTA 8460
Db 8402 ATTCAAAAGGCGAGAACTGGCGTTTACCCTGCGGTGCGCGCGAGCGCGTGTGACGACTA 8461
QY 8461 GCTGGGGTAAATACCTCATGTTTACTTTGAAGGCCACTGACGCTGTGAGCTGCAAAAGC 8520
Db 8462 GCTGGGGTAAATACCTCATGTTTACTTACCTGAAAGCCACTGCGGCCCTGTGAGCTGCAAGC 8521
QY 8521 TCCAGGCTGACAGATGTCTGTGAACGGAGACGACTTGTCTGTATCTGTGAAAGCGCGG 8580
Db 8522 TCCGGGACTGACGATGTCTGTGAACGGAGACGACTTGTCTGTATCTGTGAAGCGCGG 8581
QY 8581 GAAACCCAGGAGATCGCGCGCTTACGAGCTTTCACGGAGGCTATGACTAGTACTCTG 8640
Db 8582 GAAACCCAGGAGATCGCGCGCTTACGAGTCTTTCACGGAGGCTATGACTAGTACTCTG 8641
QY 8641 CCCCCCGGGGATCCGCCCCCAACAGAAATAGCACTGGAGCTGATAACATCATGTTCTCT 8700
Db 8642 CCCCCCGGGGACCCGCTCAACGGAAATAGCACTTGGAGTTGATAACATCATGTTCTCT 8701
QY 8701 CCAATGTGTCTGTCGCGCACGATGCTCTGGCAAAAGGGGTATACTACTCTCACCCGCTGACC 8760

8702 DB CCAATGTGTCGTCGACACGATGTCATCTGGTAAAAAGGTTGTAACCTCAACCGTGACC 8761
8761 QY CCACCAACCCCTTTGACAGGGTGGTGGAGACGTAGACACACTCAATCAACTCTTT 8820
8762 DB CTACCAACCCCTTTGACAGGGTGGTGGAGACGTAGACACACTCAATCAACTCTTT 8821
8821 QY GGCTAGGCAATATCATATGATGCGCCCAACCTATGGGCAAGGATGATTTCTGATGACTC 8880
8822 DB GGCTAGGCAATATCATATGATGCGCCCAACCTATGGGCAAGGATGATTTCTGATGACTC 8881
8881 QY ACTTTTCTCCATCTCTAGCTCAAGAGCACTTGAAGAAGCCCTGGATTTGACATCT 8940
8882 DB ATTCTTCTCCATCTCTAGCTCAAGAGCACTTGAAGAAGCCCTGGATTTGACATCT 8941
8941 QY ACGGGGCTTCTACTCATTTGAGCCACTTCACTCTAGCTCAAGAGCACTTGAAGAAGCCCTGGATTTGACATCT 9000
8942 DB ACGGGGCTTCTACTCATTTGAACCACTTGAATCTCACTCAAGATCAATGAGGACTCCATG 9001
9001 QY GTCTTAGCGCATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT 9060
9002 DB GTCTTAGCGCATTTTCACTCCATGTTACTCTCCAGGCGAGATCAATAGGGTGGCTTCAT 9061
9061 QY GCCTCAGGAAACTTTGGGTACCACTTGGAACTTGGAACTTGGAACTTGGAACTTGGAACTTGG 9120
9062 DB GCCTCAGGAAACTTTGGGTACCACTTGGAGCTTGGAGCACTTGGAGCACTTGGAGCACTTGG 9121
9121 QY GCCTTAGCTACTTGTCCAGGGGGAGGGCGGCACTTGTGGCAGATACCTCTTTAACT 9180
9122 DB GCGTAAAGCTACTTGTCCAGGGGGAGGGCGGCACTTGTGGCAGATACCTCTTTAACT 9181
9181 QY GGGCAGTAAGGACCAAGCTTAAACTCACTCACTCCAGCGCGCTCCAGCTGACATTTGT 9240
9182 DB GGGCGGTGAGGACCAAGCTTAAACTCACTCACTCCAGCGCGCTCCAGCTGACATTTGT 9241
9241 QY CTGGCTGGTTCGTGCTGTTTACAGCGGGGAGACATATATFACAGCTGTCTCGTGCCC 9300
9242 DB CCGGCTGGTTCGTGCTGTTTACAGCGGGGAGACATATATFACAGCTGTCTCGTGCCC 9301
9301 QY GACCCCGCTGTTTCCGTTGCTTACTCTCTACTCTTCTAGGGGTAGGCAATTTACCTGC 9360
9302 DB GACCCCGCTGTTTCCGTTGCTTACTCTCTACTCTTCTAGGGGTAGGCAATTTACCTGC 9361
9361 QY TCCCAACCGATGAACCGGGAGCTAACCACTCCAGGCTTAAAGCAATTTCTGTTTTTTT 9420
9362 DB TCCCAACCGATGAATGGGAGCTTAACTTCACTCCAGGCAATAGGCGCTTCTTTTTTT 9421
9421 QY TT 9471
9422 DB TT 9472

RESULT 9
US-09-539-601-1
; Sequence 1, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlagler, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)

OTHER INFORMATION: construct I389/Core-3'/wt
FEATURE:
NAME/KEY: CDS
LOCATION: (342)..(1193)
OTHER INFORMATION: hepatitis C virus core - neomycin
OTHER INFORMATION: phosphotransferase fusion protein
FEATURE:
NAME/KEY: RBS
LOCATION: (1202)..(1812)
OTHER INFORMATION: internal ribosome entry site from
OTHER INFORMATION: encephalomyocarditis virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1813)..(10845)
OTHER INFORMATION: hepatitis C virus polyprotein from core to
OTHER INFORMATION: nonstructural protein NS5B; parental sequence
OTHER INFORMATION: without cell culture-adaptive mutations
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (10846)..(11076)
US-09-539-601-1

Query Match 82.5%; Score 7918.6; DB 3; Length 11076;
Best Local Similarity 91.0%; Pred. No. 0;
Matches 8433; Conservative 0; Mismatches 824; Indels 10; Gaps 1;
QY 339 ACCATGAGCAGCAATCTTAACCTCAAGAAACCAACCAAGTAACCAACCGCGCCCA 398
DB 1810 ACCATGGGACGAACTCTAAACCTCAAGAAACCAACCAAGTAACCAACCGCGCCCA 1869
QY 399 CAGGAGCTCAAGTTCCCGGGCGGTGTGAGATGCTTGGTGGAGTTTACCTGTGGCGGC 458
DB 1870 CAGGAGCTCAAGTTCCCGGGCGGTGTGAGATGCTTGGTGGAGTTTACCTGTGGCGGC 1929
QY 459 AGGGGCGCCAGGTTGGGTGTGCGCGCACTAGGAAGGCTTCCGAGCGGTGCGCAACCTCGT 518
DB 1930 AGGGGCGCCAGGTTGGGTGTGCGCGCACTAGGAAGGCTTCCGAGCGGTGCGCAACCTCGT 1989
QY 519 GGAAGGCGCAACCTATCCCAAGGCTGCGCAACCGGAGGCGAGGCTGGGCTCAGCCC 578
DB 1990 GGAAGGCGCAACCTATCCCAAGGCTGCGCAACCGGAGGCTGGGCTCAGCCC 2049
QY 579 GGGTACCTTGGCGCTCTATGCAATAGGGCTTGGGTTGGGAGGATGGCTCTCTGCA 638
DB 2050 GGGTACCTTGGCGCTCTATGCAATAGGGCTTGGGTTGGGAGGATGGCTCTCTGCA 2109
QY 639 CCGCGCGCTCCCGGCTAGTTGGGGCCCAAGGCTGCGCAACCGGAGGCTGGGCTCAGCCC 698
DB 2110 CCGCGCTGCTCGGCTAGTTGGGGCCCAAGGCTGCGCAACCGGAGGCTGGGCTCAGCCC 2169
QY 699 GGTAAAGGTCAATCGATACCTTATACGCGGCTTCCCGATCTCATGCGGTACATTCGGCTC 758
DB 2170 GGTAAAGGTCAATCGATACCTTATACGCGGCTTCCCGATCTCATGCGGTACATTCGGCTC 2229
QY 759 GTGCGCGCCCTTAGGGGCGCTGCGAGGCGCTTGGCAGCGGTGTCGACAGGTTCTGGAG 818
DB 2230 GTGCGCGCCCTTAGGGGCGCTGCGAGGCGCTTGGCAGCGGTGTCGACAGGTTCTGGAG 2289
QY 819 GACGCGTGAATATGCAAGGAACTGCGCGGTTGCTCTTCTCTATCTCTCTCTG 878
DB 2290 GACGCGTGAATATGCAAGGAACTGCGCGGTTGCTCTTCTCTATCTCTCTCTG 2349
QY 879 GCTCTGCTGCTCTTTTGACCATCCAGCTTCCCGTTTATGAAGTGGCGCAACGTTCCGGG 938
DB 2350 GCTTGTGCTGCTGTTTGAACCATCCAGCTTCCCGTTTATGAAGTGGCGCAACGTTCCGGG 2409
QY 939 ATATACCATGTCACGAACGACTGCTCCAACTCAAGCAATTTGATGAGGAGCGGACGTG 998
DB 2410 GTGTACCATGTCACGAACGACTGCTCCAACTCAAGCAATTTGATGAGGAGCGGACATG 2469
QY 999 ATCATGCACTCTCCGGGTGCTGCTGTTTCAAGGAGGTTTACAGCTCCGTTGCTGG 1058
DB 2470 ATCATGCACTCTCCGGGTGCTGCTGTTTCAAGGAGGTTTACAGCTCCGTTGCTGG 2529

Db 4690 GCGCCAGCGGGCTACGAGACCTTGGGTGGGAGTTGAGCCCGTCTCTCTCTGATATG 4749
QY 3279 GAGACCAAGTTCATCACCTTGGGAGCAGACACCGCTGCGTGTGGGACATCATCTTTGGGT 3338
Db 4750 GAGACCAAGTTTATCACCTTGGGGGAGACACCGCGGCGGTGTGGGACATCATCTTTGGGC 4809
QY 3339 CTACCGGTCTCGCCGGAAGGGGAAGGAGATATTTTGGGACCGCTGATAGTCTCGAA 3398
Db 4810 CTGCGCGTCTCGCCGCGAGGGGAGGAGATACATCTGGGACCGGACAGACGCTTGAA 4869
QY 3399 GGGCAAGGTGGGACTCTCTGCGCCATCAGGCTCTCTCCCAACAAACGCGGGGCGTA 3458
Db 4870 GGGAGGGGTGGGACTCTCTGCGCCTATTAAGGCTCTCTCCCAACAGAGCGAGGCGCTA 4929
QY 3459 CTTGGTTGATCATCATAGCTCTACAGCGCGGGAACAAGAACAGAGTCCGAAGGGAGGTT 3518
Db 4930 CTTGGCTGATCATCACTAGCTCTACAGCGCGGACAGGAACAGAGTCCGAGGGGAGGTC 4989
QY 3519 CAAAGTGGTTTACCGCAACACATCTTCTGCGGACCTGATCAACGCGCTGTCTGG 3578
Db 4990 CAAAGTGGTCTCACCGCAACACATCTTCTTGGCGACCTGCGTCAATGGCGTGTGG 5049
QY 3579 ACTGTCTACATGGCGCTGCTCGAAGACCTAGCGGTCGCAAGAGTCCCAATCACCAA 3638
Db 5050 ACTGTCTATATGTCGCGCTCAAGACCTTGGCGGCGCAAGGGCCCAATCACCCAA 5109
QY 3639 ATGTACACCAATAGACCTTGAACCTCGTGGTGGCAGCGGCCCGCGGCGCGCTCC 3698
Db 5110 ATGTACACCAATGTGACCAAGACCTCGTGGTGGCAGCGGCCCGCGGCGCGTTC 5169
QY 3699 ATGACACCAATGAGCTGTGGAGCTCGGACCTTACTTGTGTACGAGACATGCTGATGTC 3758
Db 5170 TTGACACCAATGACCTGCGGACCTCGGACCTTACTTGTGTACGAGGCAATGCGATGTC 5229
QY 3759 ATTCGGTGGCGGCGGAGCGACAGCGGGAAGTCTACTCTCCCGCGGCGCGCTCC 3818
Db 5230 ATTCGGTGGCGGCGGAGCGACAGCGGGAAGCTACTCTCTCCCGCGGCGCGCTCC 5289
QY 3819 TACTGAAAGGCTCTCTGGGTGCTCAATGCTTGTGGCGGCGACGCTGTGGGCGTC 3878
Db 5290 TACTGAAAGGCTCTCTGGGCGGTCACTGCTCTGCGGCGGCGGCGGCGTGTGGGCGATC 5349
QY 3879 TTCGGGCTGCTGTGTGACCCCGGGGGTCCGAAGCGGTGGAATTCATACCGTTGAG 3938
Db 5350 TTTTGGGCTGCGGTGTGACCCCGAGGGTTGCGAAGCGGTGGAATTTGTACCGCTCGAG 5409
QY 3939 TCTATGAAACTACCATGCGGTCTCGGTCTTACAGAGAACTCAACCCCGCGGTGTA 3998
Db 5410 TCTATGAAACCACTATGCGGTCTCCCGGTCTTCAAGGACAACTCGTCCCGCGCGGTA 5469
QY 3999 CCGCAGACATTCGAAGTGGCACATCTGACGCTCTACTGCGAGCGGCAAGAGCACCAA 4058
Db 5470 CCGCAGACATTCAGGTGGGCGGCTCTACGCGGCTCTACTGCTAGCGGCAAGAGCACTAG 5529
QY 4059 GTGCGGCTGCTGTATGAGCGGCGGAGGTGCAAGGTGCTGCTCTGAAACCGCTCGGTGCC 4118
Db 5530 GTGCGGCTGCTGTATGAGCGGCGGAGGTGTAAGGTGCTGCTCTGAAACCGCTCGGTGCC 5589
QY 4119 GCACCTTAGGGTTGGGGCGGTATATGTCGAGGCAACCGGTATCGACCCCTAACATCAGA 4178
Db 5590 GCCACCTTAGGGTTGGGGCGGTATATGTCGAGGCAACCGGTATCGACCCCTAACATCAGA 5649
QY 4179 ACTGGGTAGGCAATTCACCGGCGGCTTACTGCTACTGCTACTGCTAGGCAAGTTC 4238
Db 5650 ACCGGGTAGGCAATTCACCGGCGGCTTACTGCTACTGCTACTGCTAGGCAAGTTC 5709
QY 4239 CTTCCGACGCTGCTGTCTGGGGGCGGCTTATGACATCATATATGATGATGAGTGGCAC 4298
Db 5710 CTTGCCAGCGGTGCTGTCTGGGGGCGGCTTATGACATCATATATGATGATGAGTGGCAC 5769
QY 4299 TCAACTGACTGACTACCATCTTGGGCGATCGGCAAGTCTCTGGACCAAGCGGAGCGGCT 4358

5770 TCAACTGACTGACCACTATCTCTGGGCACTCGGCACAGTCTCTGGACCAAGCGGAGCGGCT 5829
QY 4359 GGAGCGGCGCTGCTGCTGCTCGCCACCGCTACACCTCGGGATCGGTTACCGTGGCCACAC 4418
Db 5830 GGAGCGGCACTGCTGCTGCTCGCCACCGCTACCGCTCGGGATCGGTTACCGTGGCCACAT 5889
QY 4419 CCCAATATCGAGGAATAGGCTGTCTCAACAAATGAGAGATCCCTTCTATGCAAGGCC 4478
Db 5890 CCACATATCGAGGAGTGGCTCTGTCCAGCACTGGAGAAATCCCTTTTATGCAAGGCC 5949
QY 4479 ATCCCATTTAGGCCCATCAAGGGGGAGGCACTCTCATTTTCTGCCATTTCCAAAGAAA 4538
Db 5950 ATCCCATCGAGACCATCAAGGGGGAGGCACTCTCATTTTCTGCCATTTCCAAAGAAA 6009
QY 4539 TGTGACGAGCTCGCGCAAGCTGACAGGCTCGGACTCGGACTCAATGCTGTAGCATATTAACGG 4598
Db 6010 TGTGATGAGCTCGCGCGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGG 6069
QY 4599 GGCCTTGTATGTCGCTCATACCGCTATCGGAGAGTGTCTGCTGGCAACAGAGCT 4658
Db 6070 GGCCTTGTATGTCGCTCATACCACTAGGGAGAGCTCATTTCTGTAGCAACGGAGCT 6129
QY 4659 CTAATGACCGGTTTACCGCGGATTTGACTCAAGTATGACTGCAATATCATGCTGCTCAC 4718
Db 6130 CTAATGACCGGCTTTACCGCGATTTGACTCAAGTATGACTGCAATATCATGCTGCTCAC 6189
QY 4719 CAGACAGTGCATTTAGCTTGGATCCCATCTTACCACTTGAAGAGCAGCGGTCGCCAA 4778
Db 6190 CAGACAGTGCATTTAGCTTGGATCCCATCTTACCACTTGAAGAGCAGCGGTCGCCAA 6249
QY 4779 GACGCGTGTGCGGCTCGCAACCGGAGGTAGAACTGGCAGGGTAGGAGTGGCATCTAC 4838
Db 6250 GACGCGTGTGACGCTCGAGCGGAGGAGACTTGTAGGGGAGGATGGGCACTTAC 6309
QY 4839 AGGTTTGTGATCTCAGGAGAACCGGCTCGGCAATGTTGATTTCTCGGCTCTGTGTGAG 4898
Db 6310 AGGTTTGTGATCTCAGGAGAACCGGCTCGGCAATGTTGATTTCTCGGCTCTGTGTGAG 6369
QY 4899 TGTATGACGCGGCTGTGCTGTATGACTCAGCGGCTGAGACTCGGTTAGTGTG 4958
Db 6370 TGTATGACGCGGCTGTGCTGTATGACTCAGCGGCTGAGACTCGGTTAGTGTG 6429
QY 4959 CCGGCTTACCTTAATACACGAGGTTGCGGCTGCGCAGGACCATCTGGAGTTCTGGAG 5018
Db 6430 CCGGCTTACCTTAATACACGAGGTTGCGGCTGCGCAGGACCATCTGGAGTTCTGGAG 6489
QY 5019 AGGCTTCTTACAGGCTTACCCACATAGATGCCCCATTTCTGTCTCCAGACTTAAACAGCA 5078
Db 6490 AGGCTTCTTACAGGCTTACCCACATAGATGCCCCATTTCTGTCTCCAGACTTAAACAGCA 6549
QY 5079 GGAGACAACTTTCTTACTGTTGGCATATCAAGCTACAGTGTGCGGCGAGGCTCAAGCT 5138
Db 6550 GGAGACAACTTTCTTACTGTTGGCATATCAAGCTACAGGCTACGCTGTGCGGCGAGGCTCAGGCT 6609
QY 5139 CCACCTTCACTGGGAGCAAAATGTGAAGTGTCTCATCGGCTGAACCTTACACTGAC 5198
Db 6610 CCACCTTCACTGGGAGCAAAATGTGAAGTGTCTCATCGGCTGAACCTTACACTGAC 6669
QY 5199 GGGCAACACCCCTGCTGTATAGGCTAGGAGCGGCTCCAAATAGAGGTCATCTCTCACACAC 5258
Db 6670 GGGCAACACCCCTGCTGTATAGGCTAGGAGCGGCTTCAAAACAGAGTTACTTACACACAC 6729
QY 5259 CCCATACTAAATACATGCGCATGATGTCGCTGACCTGGAGGTCGTCACTAGCAC 5318
Db 6730 CCCATACTAAATACATGCGCATGATGTCGCTGACCTGGAGGTCGTCACTAGCAC 6789
QY 5319 TGGTGTGTTAGCGGAGTCTTGTGAGCTTGGCGGCTACTGCTGACAGAGGAGT 5378
Db 6790 TGGTGTGTTAGCGGAGTCTTGTGAGCTTGGCGGCTACTGCTGACAGAGGAGT 6849
QY 5379 GTGTCATTTGGGAGGATCATCTTGTCCGGGAAGCGGAGTGTGCTTCCCGACAGGGA 5438
Db 6850 GTGTCATTTGGGAGGATCATCTTGTCCGGGAAGCGGAGTGTGCTTCCCGACAGGGA 6909

Db 9070 TCGATGTCCTACATGAGCAGGCGCCCTGATACGCCATGCGCTGCGGAGGAAACCAAG 9129
Qy 7659 CTGCCATCAACCGTTGAGCAACTCTTTTGTGCGTCAACACAAACATGCTGCTACGCCACA 7718
Db 9130 CTGCCATCAATGACACTGAGCAACTCTTTGCTCGTCAACACAACTTTGGTCTATGCTACA 9189
Qy 7719 ACATCCGCGAGCGCAAGCCTCCGCGCAGAGAGGTCACTTTTGACAGATTGCAAGTCTCG 7778
Db 9190 ACATCTCGCAGCGCAAGCCTCCGCGCAGAGAGGTCACTTTTGACAGACTGCAAGTCTCG 9249
Qy 7779 GATGATCATTTACCGGACGTACTCAAGGAGATGAAGCGAGCGTCCACAGTTTAAGGCT 7838
Db 9250 GACGACCACTACCGGAGCGTCTCAAGGAGATGAAGCGAGCGTCCACAGTTTAAGGCT 9309
Qy 7839 AAGCTTCTATCTATGAGGAGCGCTGCAAGCTGACGCGCCCAACATTCGCGCAATCCAAA 7898
Db 9310 AACTTCTATCCGTGAGGAGCGCTGTAAGCTGACGCGCCCAACATTCGCGCAATCCAAA 9369
Qy 7899 TTTGGCTATGGGCAAGGAGCGTCCGGAACCTATCCAGCAGGCGCTTTAACCAATCCGC 7958
Db 9370 TTTGGCTATGGGCAAGGAGCGTCCGGAACCTATCCAGCAGGCGCTTTAACCAATCCGC 9429
Qy 7959 TCGTGTGGAGGACTTGTCTGGAAGACACTGAAACACCAATTTGACACCACTCATGGCA 8018
Db 9430 TCGTGTGGAAGGACTTGTCTGGAAGACACTGGAACACCAATTTGACACCACTCATGGCA 9489
Qy 8019 AAAAGTGAAGTTTCTGCTGTCACAGAGAGGAGCGCGCAAGCAGCTTCGCTTTATC 8078
Db 9490 AAAAATGAGTTTCTGCTGTCACAGAGAGGAGCGCGCAAGCAGCTTCGCTTTATC 9549
Qy 8079 GTATTCCAGACCTGGGAGTTGCTGTATGCGAGAGATGGCCCTTTACGAGTGGTCTCC 8138
Db 9550 GTATTCCAGATTGGGGTTGCTGTGTGCGAGAAATGGCCCTTTACGATGTGCTCC 9609
Qy 8139 ACCCTTCTCAGGCGGTGATGGCTCTCATACGGAATTTCAATACTCCCCCAAGCAGCGG 8198
Db 9610 ACCCTTCTCAGGCGGTGATGGCTCTCATACGGAATTTCAATACTCCCCCAAGCAGCGG 9669
Qy 8199 GTCCAGTCTCGTGTGATACCTGTGAATCAAGAAATGSCCTATGGGCTTCTCATATGAC 8258
Db 9670 GTCCAGTCTCGTGTGATACCTGTGAATCAAGAAATGSCCTATGGGCTTCTCATATGAC 9729
Qy 8259 ACCCGCTGTTTGTGACTCAACCGTCACTGAGAGTGACATTCGTGTGAGGAGTCAATTTAC 8318
Db 9730 ACCCGCTGTTTGTGACTCAACCGTCACTGAGAAATGACATCCGTGTGAGGAGTCAATCTAC 9789
Qy 8319 CAATGTTGTGACTTGGCCCCCGAGGCCAGACAGGCCATAAGGTCTGCACAGAGCGGCTT 8378
Db 9790 CAATGTTGTGACTTGGCCCCCGAGGCCAGACAGGCCATAAGGTCTGCACAGAGCGGCTT 9849
Qy 8379 TACATCGGGGTCCCTGACTAACTCAAAAGGCGAGACTCGGTTATCGCGGTGCGC 8438
Db 9850 TACATCGGGGTCCCTGACTAACTCAAAAGGCGAGAACTCGCGGTATCGCGGTGCGC 9909
Qy 8439 GCAAGTGCCTGTGACGACTAGCTGCGGTAACTACCTCAATGTTACTTGAAGGCCACT 8498
Db 9910 GCGAGCGGTGACTGACCAACGACTGCGGTAACTACCTCAATGTTACTTGAAGGCCCT 9969
Qy 8499 GCAGCCTGTGAGCTGCAAGCTCCAGGACTGCAGGATGCTCGTGAACCGGAGACGACCTT 8558
Db 9970 GCGGCTGTGAGCTGCGAGCTCCAGGACTGCAGGATGCTCGTATGCGGAGACGACCTT 10029
Qy 8559 GTGCTTATCTGTGAAGCGCGGAAACCCAGAGGAGATGCGGCGCCCTACGAGCCTTACG 8618
Db 10030 GTGCTTATCTGTGAAGCGCGGAAACCCAGAGGAGAGCTACGAGCCTTACG 10089
Qy 8619 GAGGCTATGACTAGTATTCCGCCCGCCCGGATCCGCCCAACCAACAGATACGACCTG 8678
Db 10090 GAGGCTATGACTAGTATTCTGCCCGCCCGGATCCGCCCAACCAACAGATACGACTG 10149
Qy 8679 GAGCTGATAACATCATGTTCTCTCAATGTGTGCTGCGGACGATGCTATGCGCAAAAGG 8738

Db 10150 GAGTGTATTAACATCATGCTCTCCAATGTGTGAGTGGCGCAGCATGCTGCGCAAAAGG 10209
Qy 8739 GTATATCTACCTACCGGTGACCCCAACACCCCTTTCGACGGGCTGGTGGGAGACAGCT 8798
Db 10210 GTGTACTATCTACCGGTGACCCCAACACCCCTTTCGCGGGCTGGTGGGAGACAGCT 10269
Qy 8799 AGACACACTCCAATCAACTCTTGGCTAGGCAATATCATGATATGCGCCACCACTATGG 8858
Db 10270 AGACACACTCCAGTCAATCTTGGCTAGGCAACATCATGATATGCGCCACCACTTGG 10329
Qy 8859 GCAAGATGATTTCTGATGACTCACTTTTCTCCATCTTCTAGCTCAAGACAACTTGAA 8918
Db 10330 GCAAGATGATCTGATGACTCACTTTTCTCCATCTTCTAGCTCAGGAACAACCTGAA 10389
Qy 8919 AAAGCCTTGAATTTGTAGATCTACGCGGCTTGTACTCCATTTAGGCCACTTGACTACCT 8978
Db 10390 AAAGCCTTAGATTTGTAGATCTACGCGGCTTGTACTCCATTTAGGCCACTTGACTACCT 10449
Qy 8979 CAGATCAATCAAGACTCTTCAATGCTTGTAGCGCATTTACACTCCACAGTTACTCTCCAGGT 9038
Db 10450 CAGATCAATCAAGACTCTTCAATGCTTGTAGCGCATTTTCACTCCATAGTTACTCTCCAGGT 10509
Qy 9039 GAGATCAATAGGTGGCTTTCATGCTCAGGAACTTGGGGTACCACTTTCGCAACCTGG 9098
Db 10510 GAGATCAATAGGTGGCTTTCATGCTCAGGAACTTGGGGTACCGCTTTCGAGTCTGG 10569
Qy 9099 AGACATCGGCGCAGAGTGTCCGCTTAAGCTACTGTCCAGGGGGGAGGGCGCACT 9158
Db 10570 AGACATCGGCGCAGAGTGTCCGCTTAAGCTACTGTCCAGGGGGGAGGGCTGCCACT 10629
Qy 9159 TGTGGAGATACCTCTTTAATCTGGGAGTAAGGACCAAGCTTAACTCACTCCCAATCCCG 9218
Db 10630 TGTGGCAAGTACCTCTTTCACTGGGAGTAAGGACCAAGCTCAAACTCACTCCCAATCCCG 10689
Qy 9219 GCGCGCTCCAGCTGGAGTGTGTGGCTGGTTCGTGCTGCTTACAGCGGGGAGACATA 9278
Db 10690 GCTGCGTCCAGTTGGATTTATCCAGCTGGTTCGTGCTTACAGCGGGGAGACATA 10749
Qy 9279 TATCACAGCTGTCTGTCGCCGACCCCGCTGGTTCCGTTGTCCTACTCTACTTTCT 9338
Db 10750 TATCACAGCTGTCTGTCGCCGACCCCGCTGGTTTCATGCTGCTACTCTACTTTCT 10809
Qy 9339 GTAGGGTAGGATTTACTGCTCCCAACCGATGAAGGGGAGCTAACCACTCCAGGCC 9398
Db 10810 GTAGGGTAGGATTTACTTACTCCCAACCGATGAAGGGGAGCTAACCACTCCAGGCC 10869
Qy 9399 TTAAGCC-----ATTTCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9448
Db 10870 AATAGGCAATCCTGTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 10929
Qy 9449 TTTTCTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 9508
Db 10930 TTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 10989
Qy 9509 CTTAGCCCTAGTCAAGCTAGCTGTGAAGTCCGTCGTCGAGCCGATGACTGAGAGAGTGC 9568
Db 10990 CTTAGCCCTAGTCAAGCTAGCTGTGAAGTCCGTCGTCGAGCCGCTTGTACTGAGAGAGTGC 11049
Qy 9569 TGATCTGGCCTCTCTGAGATCATGT 9595
Db 11050 TGATCTGGCCTCTCTGAGATCAAGT 11076

RESULT 10

US-09-539-601-25
; Sequence 25, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlagher, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30

```

; EARLIER APPLICATION NUMBER: 199 175 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/5.1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neo
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein
; OTHER INFORMATION: nonstructural protein NS5B;
; OTHER INFORMATION: culture-adaptive mutations of
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
; US-09-539-601-25

```

Query Match.	82.3%;	Score 7901;	DB 3;	Length 11076;
Best Local Similarity	90.9%;	Pred. Mis. 0;		
Matches 8422;	Conservative	0;	Matches 835;	Indels 10; Gaps 1;
Qy	339	ACCATGACGACGAACTCTAAACCTCAAGAAACCAACAAACGTAACCAACACCGCCGCCCA	398	
Db	1810	ACCATGGCAGCAATCTTAAACCTCAAGAAACCAACAAACGTAACCAACACCGCCGCCCA	1869	
Qy	399	CAGGACGCTCAAGTTTCCCGGGCGGTGGTTCAGATCGTTGGTGGAGTTTACCTGTTGCCCGCGC	458	
Db	1870	CAGGACGCTCAAGTTTCCCGGGCGGTGGTTCAGATCGTTCGGTGGAGTTTACCTGTTGCCCGCGC	1929	
Qy	459	AGGGGCCCCAGGTTGGGTGTGCGCGCGACTAGGAAGGCTTCGAGCGGTTCGCAACCTCGT	518	
Db	1930	AGGGGCCCCAGGTTGGGTGTGCGCGCGACTAGGAAGACTTCCGAGCGGTTCGCAACCTCGT	1989	
Qy	519	GGAGGGCGCAACCTATCCCAAGGCTCGCGACCCGAGGCGAGGGGCTTGGGCTCAGGCC	578	
Db	1990	GGAGGGCGCAACCTATCCCCAAGGCTCGCCAGCCGAGGCTGAGGGCTTGGGCTCAGGCC	2049	
Qy	579	GGGTACCCCTTGGCCCTCTATGGCAATGAGGGCTTGGGGTGGGAGGATGGCTCTCTGTCA	638	
Db	2050	GGGTACCCCTTGGCCCTCTATGGCAATGAGGGCTTGGGGTGGGAGGATGGCTCTCTGTCA	2109	
Qy	639	CCCCCGGGCTCCGGGCTAGTTGGGGGCCCAACGACCCCGCGCTTAGGTTCGCGTAATTG	698	
Db	2110	CCCCGTGGCTCTCGGGCTAGTTGGGGGCCCAACGACCCCGCGCTTAGGTTCGCGCAATTTG	2169	
Qy	699	GGTAAGGTCATGATACCTTACATGGGCTTCCCGCATCTCATGGGGTACATTCGCGTC	758	
Db	2170	GGTAAGGTCATCGATAACCTTACATGGGCTTCCCGCATCTCATGGGGTACATTCGCGTC	2229	
Qy	759	GTCCGGCCCCCTTAGGGGGCGCTGCCAGGGCTTTGGCACACGGTGTCCGGGTTCTGGAG	818	
Db	2230	GTCCGGCCCCCTTAGGGGGCGCTGCCAGGGCTTGGCGCATGCGGTCTCGGATCTGGAG	2289	
Qy	819	GACGGCGTGAACATATGCAACAGGGAACTTGGCCCGGTTGCTCTTTCTCTATCTTCCTCTTG	878	
Db	2290	GACGGCGTGAACATATGCAACAGGGAACTTGGCCCGGTTGCTCTTTTCTATCTTCCTTTTG	2349	

Qy	879	GCTCTGCTCTCTGTTTGACCATCCAGCTTCGGCTTATGAAGTGGCCAACTGTCGGGG	938
Db	2350	GCTTTGCTGCTCTGTTTGACCATCCAGCTTCGGCTTATGAAGTGGCCAACTGTCGGGA	2409
Qy	939	ATATACCATGTACAGAACGACTGCTCAAATCAAGCAATTGTGTATGAGGCAGCGGACGTG	998
Db	2410	GTGTACCATGTACAGAACGACTGCTCAAAGCAAGCATTTGTGTATGAGGCAGCGGACATG	2469
Qy	999	ATCATGCATACTCCGGGTGCGTSCCTGTGTATTACGAGGGTTAAACAGTCCCGTTGCTGG	1058
Db	2470	ATCATGCATACCCCGGTGCGTGCCCTCGTTTCGGGAGAAACAACTCTCCCGCTGCTGG	2529
Qy	1059	GTAGCGCTACTCCGACGCTCGCGGCAGGAATGCCAGAGTCCCGACTACGACAATAACGA	1118
Db	2530	GTAGCGCTACTCCGACGCTCGCGGCAGGAACGCTAGCGTCCCGCACTACGACGATACGA	2589
Qy	1119	CGCACGTGCAGTTGCTCGTTGGGACGGCTTTCTGCTCCGCTATGTACGTGGGGGAT	1178
Db	2590	CGCCATGTCCATTTGCTCGTTGGGGGGCTGCTCTCTGCTCCGCTATGTACGTGGGAGAT	2649
Qy	1179	CTCTGGCGATCTATTTTCTCGTCTCCAGCTGTTCACTTTCTCGCTCGCCGGCATGAG	1238
Db	2650	CTCTGGCGATCTGTTTCTCGTGGCCAGCTGTTCACTTTCTCGCTCGCCGGCAGAG	2709
Qy	1239	ACAGTGAGGACTGCAACTGCTCAATCTATCCCGGCATGTATCAGGTCAACCGATGGCT	1298
Db	2710	ACAGTACAGGACTGCAATTTGCTCAATATATCCCGGCACGCTGACAGGTCAACCGTATGGCT	2769
Qy	1299	TGGGATATGATGATGACTGGTCACTACAAAGCCCTAGTGTGTGCGAGTTGCTGCGGGGCTT	1358
Db	2770	TGGGATATGATGATGAACTGGTCACTTACAGCAGCCCTAGTGTGTGCGAGTTACTTCCGG	2829
Qy	1359	ATCCCAAGCTGTCGTGGACATGTGGCGGGGGCCACCTGGGGAGTCTCTGGCGGGGCTT	1418
Db	2830	ATCCCAAGCTGTCGTGGATATGTGTGGCGGGGCCCATGTGGGAGTCTCTAGCGGGCTT	2889
Qy	1419	GCCTACTATTCCATGTAGGAACTGGGCTAAGGTTCTGATGTGGCGCTACTCTTTGGC	1478
Db	2890	GCCTACTATTCCATGTGGGGAATGGGCTAAGGTTCTGATGTGTGATGCTACTCTTTGCC	2949
Qy	1479	GGCGTTGACGGGGAGCCACACGACGGGAGGTGGCGGCCACACCACTCGGGTTC	1538
Db	2950	GGCGTTGACGGGGGAACCTATGTACAGGGGGGACAGATGGCCAAAACACCCCTCGGGATT	3009
Qy	1539	ACGTCCCTTTTCTCATCTGGGGGCTCTCAGAAAAATCCAGCTTGTGAATACCAACGGCAGC	1598
Db	3010	ACGTCCCTCTTTTCCACCGGGTCAATCCAGAAAAATCCAGCTTGTGAATACCAACGGCAGC	3069
Qy	1599	TGGCACATCAACAGGACTGCCCTAAATTTGCAATGACTCCTCCAACTGGGTTCTTTGGC	1658
Db	3070	TGGCACATCAACAGGACTGCCCTGAACTGCAATGACTCCTCCCTCAACCTGGGTTCTTTGCT	3129
Qy	1659	GGCGTGTTTTACGCACACAAGTTTCAACTGCTCGGGTGCCGGAGCGCATGGCCAGCTGC	1718
Db	3130	GGCGTGTTTTACGTGCAACAAGTTTCAACTCATCTGGAATGCCAGAGCGCATGGCCAGCTGC	3189
Qy	1719	CGCCCCATTGACTGTGTTTCGCGAGGGGTGGGGGCCCATCACTATATAAGACCTTAAACAGC	1778
Db	3190	AGCCCCATTGACGGGTTCCGTCAGGGGTGGGGGCCCATCACTTACATGAGTCACACAGC	3249
Qy	1779	TCGGATCAGAGGCCCTTATTCTGGCATTTACGCGCTCGACCGTGTGGTGTCTGTAACCCGCG	1838
Db	3250	TCGGACCAAGGCCCTTATTGTTGGCATCTACGACCCCGCGCGGTATCGTACCCGCG	3309
Qy	1839	TCGCAAGTGTGTGTCAGTGTATTGTTTTCACCCCAAGCCCTGTTGTGTGGGAGCACACC	1898
Db	3310	CGCGAGGTGTGTGTCTCAGTGTACTGCTTTCACCCCAAGCCCTGTGCTGTGTGGGAGCAC	3369
Qy	1899	GATCGTTCCCGTGTCCCTACTGATAGCTGGGGGAGAAATGAGACAGACGTGATGCTCTCTC	1958
Db	3370	GACCGGTTCCGCGTCCCTACTGATCAGTTGGGGGAGAAATGAGACGACGTGCTGCTTCTT	3429
Qy	1959	AACAACACGCGTCCGCCACAAGGCAACTGTTTCGGCTGTACATGGATGAATAGTACTGGG	2018

[illegible]

Db	4510	ATGGTGCTCCAGGCTGGTATAACAAAGTGCGGTACTTCTGTCGGCGCACACGGGCTCAAT	4569
Qy	3099	CGTGCAATGCATTTAGTGCAGAAAGTCGCGGGGTCAATTATGTCCAAATGGTCTTTCATG	3158
Db	4570	CGTGCAATGCATCTGGTCGGGAAGGTTGTCTGGGGGTCAATTATGTCCAAATGGCTCTCATG	4629
Qy	3159	AAGCTGGGCGGCTGACAGGTACGTATTATAACCACTCTTACCCTACTGCGGACTGG	3218
Db	4630	AAGTTGGCGCACTGACAGGTACGTATTATGACCACTCTCACCCACTGCGGACTGG	4689
Qy	3219	GCCACGCGGGCTACGAGACCTTTCGGGTGGCGGTAGAGCCCGTCTGTCTTCTCCGCCATG	3278
Db	4690	GCCACGCGGGCTACGAGACCTTTCGGGTGGCAGTTGAGCCCGTCTGTCTTCTCTGATATG	4749
Qy	3279	GAGACCAAGGTCAATCACTGGGGAGCAGACACCGCTGGGTGGGGACATCATCTTGGGT	3338
Db	4750	GAGACCAAGGTTATCACTTGGGGGAGCAGACCGCGGCGTGTGGGGACATCATCTTGGGC	4809
Qy	3339	CTACCCGCTCCGCGCGAAGGGGAGGAGATATTATTTGGGACCGCTGATAGTCTCGAA	3398
Db	4810	CTGCGCGTCTCGCGCCGCGAGGGGAGGGAGATACATCTGGGACCGGACAGACCTTGAA	4869
Qy	3399	GGCAAGGGTGGCGACTCTCTTTCGGCCCATACGGCCTACTCCCAACAAACGCGGGCGTA	3458
Db	4870	GGCAGGGGTGGCGACTCTCTCGGCCCTATTACGGCCTACTCTCCACAGACGCGAGCCTA	4929
Qy	3459	CTTGGTTGCATCAATCACTAGCCTCAAGCGCGGGAACAAGAACAGGTCCAAAGGGAGGTT	3518
Db	4930	CTTGGTGCATCAATCACTAGCCTCAAGCGCGGGAACAGGAACAGGTTCGAGGGGAGGTC	4989
Qy	3519	CAAGTGGTTCTACCGCAACAACATCTTTCCTGGCGACCTGCATCAACGGCTGTCTCG	3578
Db	4990	CAAGTGGTCTCACCGCAACAACATCTTTCCTGGCGACCTGCATCAATGGCGTGTGTGG	5049
Qy	3579	ACTGTCTACCATGGCGCTGGCTCGAAGACCTAGCCGGTCCAAAAGGTCCAATCACCCAA	3638
Db	5050	ACTGTCTATCATGGTGC GGCTCAAGACCCCTTGC CGGCCCAAGGGGCCCAATCACCCAA	5109
Qy	3639	ATGTACACCAATGACACTTGGACCTTCGTGGGTGGCAGGCGCCCCCGGGGCGCGCTCC	3698
Db	5110	ATGTACACCAATTGGACAGGACCTCGTCGGCTGGCAAGCGCCCCCGGGGCGCGTCC	5169
Qy	3699	ATGACACCATGACGTGTGCGAGCTCGGACCTTTACTTGGTCACAGAGACATGCTGATGTC	3758
Db	5170	TTGACACCATGACACTCGCGCAGCTCGGACCTTTACTTGGTTCAGAGGACATGCCGATGTC	5229
Qy	3759	ATTCCGGTGGCGCGGAGCGACAGCAGGGGAAGTCTACTCTCTCCCCAGGCCGCTCTCC	3818
Db	5230	ATTCCGGTGGCGCGGGGCGACAGCAGGGGAGCCTACTCTCTCCCGAGGCCGCTCTCC	5289
Qy	3819	TACCTGAAGGCTCTCTGGGTGGTCCATGCTTTGGCTTTCGGGGACGCTGTGGGGCTC	3878
Db	5290	TACTTGAAGGCTCTTCGGCGGTCCACTGCTCTGCCCCCTCGGGGACGCTGTGGGCATC	5349
Qy	3879	TTCCGGGCTGTGTGTCACCCGGGGGTTCGGAAGCGCGGTGACATCTATACCCGTTGAG	3938
Db	5350	TTTCCGGGCTGTGCTGTGTCACCCGAGGGGTTGCGAAGCGCGGTGACATTTGTACCCGTCGAG	5409
Qy	3939	TCTATGAAATACCATGCGGTCTCGGCTCTTCAACAGACAATCTCAACCCCCCGGGGTGTA	3998
Db	5410	TCTATGAAACCATATGCGGTCTCCCGTCTTCAACGGAACACTCGTCTCCCTCGGCCGTA	5469
Qy	3999	CCGAGACATTTCCAAGTGGACATCTGACCGCTCTTCTGGCAGCGGCAAGAGCACCAAA	4058
Db	5470	CCGAGACATTTCCAGGTGGCCCATCTACACGCCCTACTGTTAGCGGCAAGAGCACTAAG	5529
Qy	4059	GTGCGGGTGGTATGAGCCCAAGGTTACAAGGTGCTCTGCTGAACCCGTCGCTGCC	4118
Db	5530	GTGCGGGTGGGTATGACAGCCCAAGGGTATAAGGTGCTTCTCTGAAACCGCTCCGTCGCC	5589
Qy	4119	GCCACCTTAGGTTTGGGGCGTATATGTCGAAGGCACACGGTATCGACCTTAACATCAGA	4178
Db	5590	GCCACCTTAGGTTTGGGGCGGTATATGTCGTAAGGCACATGGTATCGACCTTAACATCAGA	5649

Db	7810	TTGCGCGGATGTCGGGGAGTCCCTCTCTCTCATGTCAACGTGGGTACAAGGGAGTCTGG	7869
Qy	6399	CGGGGGACGGCATCATGCAAAACCACTGCCCATCGGGAGCACAGATCGCGGACATGTC	6458
Db	7870	CGGGGGACGGCATCATGCAAAACCACTGCCCATGTGGGGACACAGATCACGGACATGTG	7929
Qy	6459	AAAAACGGTTCCATGAGGATCGTAGGGCTAGAACCTGCAGCAACACGTTGGCAGCAACG	6518
Db	7930	AAAAACGGTTCCATGAGGATCGTTGGGGCTTAGGACCTGTAGTAACACAGTGGCATGGAACA	7989
Qy	6519	TTCCCCATCAACCGCATACACACCGGACCTTGACACACCTCCCGCGCGCCAACTATTTC	6578
Db	7990	TTCCCATTAACCGCATACACACCGCGCCCTGCACGCCCTCCCGCGCGCAAAATTATTCT	8049
Qy	6579	AGGGCGCTATGGCGGCTGCTGAGGAGTAGTGGAGGTTACGGCTGTGGGGGATTTTC	6638
Db	8050	AGGGCGCTGTGGCGGCTGCTGAGGAGTAGTGGAGGTTACGGCGGTGGGGATTTTC	8109
Qy	6639	CACCTAGTGACGGGCATGACCACTGACAAAGTAAAGTGCCCATGCCAGGTTCCGGCCCC	6698
Db	8110	CACCTAGTGACGGGCATGACCACTGACAGCGTAAAGTGCCCGTGTGAGTTCCGGCCCC	8169
Qy	6699	GAATTTCTACGGAGGTGGATGGAGTGGTTGCAACAGGTACGCTCCGGCGTGCAAACT	6758
Db	8170	GAATTTCTACAGAAAGTGGATGGGTTGGGTTGTCACAGGTACGCTCCAGCGTGCAAACCC	8229
Qy	6759	CTTCTACGGGAGGACGTCAGTTCCAGTTCGGGCTCAACCAATACTTGGTCGGGTGCGAG	6818
Db	8230	CTCCTACGGGAGGAGTGCATTTCTGGTTCGGGCTCAATCAATACTTGGTTGGGTACAG	8289
Qy	6819	CTCCATGCGAGCCCGAACCGGACGTAAACAGTGTCTTACTTCCATGTCTACCCGATCCCTCC	6878
Db	8290	CTCCATGCGAGCCCGAACCGGATGTAGCAGTGTCTCACTTCCATGTCTACCGACCCCTCC	8349
Qy	6879	CACATTAACAGACAGACGGCTAAGGTAGGCTGGCTAGAGGTTCTCCGCCCTTTTAGCC	6938
Db	8350	CACATTAACGGCGGAGACGGCTAAGCGTAGGCTGGCCAGGGGATCTCCTCCGCCCTTTGGCC	8409
Qy	6939	AGCTCATCAGTAGCAGTGTCTGCGCTTCTTTGAAGCGACATGCATACCCACCAT	6998
Db	8410	AGCTCATCAGTAGCAGTGTCTGCGCTTCTTTGAAGCGAACATGCATACCCGTCAT	8469
Qy	6999	GACTCCCGGACGTGACCTCATCGAGGCCAAACCTCTTGTGGCGGACGAGATGGCGGA	7058
Db	8470	GACTCCCGGACGTGACCTCATCGAGGCCAAACCTCTGTTGGCGGACGAGATGGCGGG	8529
Qy	7059	AACATCATCTCGCTGGAGTCAGAGATTAAGTAGTAATTTCTGAGCTTTTCGAAACGCTT	7118
Db	8530	AACATCACCCGCTGGAGTCAGAGATTAAGTAGTAATTTTGGACTCTTTTCGAGCCGCTC	8589
Qy	7119	CACGGGAGGGATGAGAGGAGATATCGTCCGCGCGGAGATCTCTCGCAAAATCCAGG	7178
Db	8590	CAACGGAGGAGATGAGAGGAGATATCGGTTCCGGCGGAGATCTCTCGGAGGTCCAGG	8649
Qy	7179	AAGTTCCCTCAGCGTTGCCCATATGGGACGCGCGGACTACAATCTCCACTGCTAGAG	7238
Db	8650	AAATTTCCCTCGAGCGATGCCATATGGGACGCGCGGATTAACAACCTCCACTGTTAGAG	8709
Qy	7239	TCCTGGAAGGACCGGACTACGTCCTCCGTGGTACACGGATGCCCATTTGCCACCTACC	7298
Db	8710	TCCTGGAAGGACCGGACTACGTCCTCCGTGGTACACGGGTGTCCATTTGCCGCTGCC	8769
Qy	7299	AAGGCTCTCCAAATACCACTCTCCAGGAGAAAGGAGGTTGTCCTGACAGAAATCCAAT	7358
Db	8770	AAGGCCCTTCGATACCACTCTACGGAGGAGGAGCGTTGTCTGTTCAGAAATCTACC	8829
Qy	7359	GTGCTCTCTCGGCGAGCTCGCCACTAAGACCTTTCCGTAGCTCCGGATCGTCGGCC	7418
Db	8830	GTGCTCTCTCGGCGAGCTCGCCACAGAGACCTTCGGCAGCTCCGAAATCGTCGGCC	8889
Qy	7419	GTGTATAGCGGACCGGACCGGCCCTTCTGACTGGCCCTCCGACGACGAGTGAACAAGGA	7478

Qy	8559	GT	CGTTATCTGTGAAGCGGGGAACCCAGAGAGATGCGCGGCCCTTACGAGACCTTCAAG	8618
Db	10030	GT	CGTTATCTGTGAAGCGGGGACCCAAGAGAGAGCGAGGAGCTTACGGGCTTCAAG	10089
Qy	8619	GAG	CTATGACTAGGTATTTCCGCGCCCCCGGGATCCGCGCCAAACAGAAATACGACTG	8678
Db	10090	GAG	CTATGACTAGATATCTTGC GCCCTCTGGGACCCGCCCAAACCAAGAAATAGACTTG	10149
Qy	8679	GAG	CTGATAAATCATATGTTTCCTCCAAATGTTCAGTGGCGACGATGCACTTGGCAAAAG	8738
Db	10150	GAG	TGATAACATCATGCTCTCTCAATGTGTCAGTGGCGACGATGCTATGGCAAAAG	10209
Qy	8739	GT	ATACTACTCCCGTGCACCCACACCCCTTGCACGGCTGCGTGGGAGACAGCT	8798
Db	10210	GT	GTACTATCTACCCGTTGACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCT	10269
Qy	8799	AG	ACACACTCCAATCAACTCTTGGCTAGGCAATATCATCATGTATGCGCCCAACCTATGG	8858
Db	10270	AG	ACACACTCCAGTCAATTCCTGGCTAGGCAACATCATGTATGCGCCCAACCTTGG	10329
Qy	8859	GCA	AGGATGATCTGATGACTCATCTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTGAA	8918
Db	10330	GCA	AGGATGATCTGATGACTCATCTTTCTCCATCCTTCTAGCTCAGGAAACAATTGAA	10389
Qy	8919	AA	AGCCCTCGATTTGATGATCTACAGGCTTGCATCTCCATTTAGCCACTTTGACCTACT	8978
Db	10390	AA	AGCCCTCGATTTGATGATCTACAGGCTTGTACTCTCATTTAGCCACTTTGACCTACT	10449
Qy	8979	CAG	ATCATTGAACGACTCCATGGTCTTAGCGATTTTACATCTCCACAGTTACTTCCAGGT	9038
Db	10450	CAG	ATCATTCAACGACTCCATGGCTTAGCGATTTTCACTCCATAGTTACTTCCAGGT	10509
Qy	9039	GAG	ATCAATAGGCTGGCTTCATGCTCAGGAACTTGGGTACACCTTCGGACCTGG	9098
Db	10510	GAG	ATCAATAGGCTGGCTTCATGCCTCAGGAACTTGGGGTACCCTTCGGAGTCTGG	10569
Qy	9099	AG	ACATCGGCGCCAGAAGTGTCCGCTTAAGCTACTGTCCAGGGGGGAGCGCCCACT	9158
Db	10570	AG	ACATCGGCGCAGAGTGTCCGCTTAGCTACTGTCCCAGGGGGGAGGGCTGCCACT	10629
Qy	9159	TG	TGGCAGATACCTCTTTTAATCTGGGAGTAAGGACCAAGCTTAAACTCATCTCCAATCCG	9218
Db	10630	TG	TGGCAAGTACCTCTTCAACTGGGAGTAAGGACCAAGCTCAAACTCACTCCAATCCG	10689
Qy	9219	GCC	CGTCCAGCTGGAGTTGTCTGCTGGTCTGCTCGCTGTTTACAGCGGGGAGACATA	9278
Db	10690	GTC	CGTCCAGTTGGATTTATCAAGTGGTTCGTTGCTTGGTTTACAGCGGGGAGACATA	10749
Qy	9279	TAT	CACAGCCTGTCTGTGCGCCAGCCCGCTGGTTTCCGTTGTGCTACTCTCTACTTTCT	9338
Db	10750	TAT	CACAGCCTGTCTGTGCGCCAGCCCGCTGGTTTCCATGTGGTGTCTACTCTCTACTTTCT	10809
Qy	9339	GTA	GGGTAGGCATTATCTGCTCCCAACCGATGAACGGGAGCTTACCACTCCAGGCC	9398
Db	10810	GTA	GGGTAGGCATCTATCTACTCCCCAACCGATGAACGGGAGCTTAAACACTCCAGGCC	10869
Qy	9399	TTA	AGCC-----ATTTCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	9448
Db	10870	AA	TAGGCCATCTGTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	10929
Qy	9449	TTTT	TTCTTCTCTTCTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT	9508
Db	10930	TTTT	TTTTTTTTTCTCTCTTTTTTTTTTCTCTTTTTTTTTTCTTTTCTTTTCTTTTCTTTTCT	10989
Qy	9509	CTT	AGCCCTAGTCAAGCTAGCTGTGAAGGTCCTGGTAGCCGATGACTGACGAGAGTGC	9568
Db	10990	CTT	AGCCCTAGTCAAGGCTAGCTGTGAAGGTCCTGGTAGCCGCTGACTGACGAGAGTGC	11049
Qy	9569	TGA	TACTGGCTCTCTGCAGATCATGT	9595
Db	11050	TGA	TACTGGCTCTCTGCAGATCAAGT	11076

```

RESULT 11
US-09-539-601-19
; Sequence 19, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURES:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/9-13F
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein from core to
; OTHER INFORMATION: nonstructural protein NS5B: carries cell
; OTHER INFORMATION: culture-adaptive mutations from clone 9-13F
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
;
US-09-539-601-19

Query Match      82.3%; Score 7899.4; DB 3; Length 11076;
Best Local Similarity 50.9%; Pred. No. 0;
Matches 8421; Conservative 0; Mismatches 836; Indels 10; Gaps 1;

QY 339 ACCATGACGACGAATCCTAAACCTCAAAGAAAAACCAACGTAACCAACCGCGGCCCA 398
DB 1810 ACCATGGGACGAATCCTAAACCTCAAAGAAAAACCAACGTAACCAACCGCGGCCCA 1869
QY 399 CAGGACGTCAAGTTCCCGGCGGTGGTGCAGATCGTTGGTGAGTTTACCTGTTGCCGCGC 458
DB 1870 CAGGACGTCAAGTTCCCGGCGGTGGTGCAGATCGTGGTGAGTTTACCTGTTGCCGCGC 1929
QY 459 AGGGGCCCCAGTTGGGTGTGCGGCGCACTAGGAAGCTTCCGAGCGGTGCGAACCTCGT 518
DB 1930 AGGGGCCCCAGTTGGGTGTGCGGCGCACTAGGAAGACTTCCGAGCGGTGCGAACCTCGT 1989
QY 519 GGAAGGCGACAACCTATCCCAAGGCTCGCGCACCCGAGGCGAGGGCTGGGGCTCAGCCC 578
DB 1990 GGAAGGCGACAACCTATCCCAAGGCTCGCGCACCCGAGGCGAGGGCTGGGGCTCAGCCC 2049
QY 579 GGGTACCTTTGGCCCTCTATGCGAATGAGGGCTCTGGGGTGGGACGATGGCTCTGTCA 638
DB 2050 GGGTACCTTTGGCCCTCTATGCGAATGAGGGCTTGGGGTGGGACGATGGCTCTGTCA 2109
QY 639 CCCCGGGCTCCGGCTAGTTGGGGCCCCACGAGACCCCGGCGTAGGTGCGTAACCTTG 698
DB 2110 CCCCGGGCTCTGGGCTAGTTGGGGCCCCACGAGACCCCGGCGTAGGTGCGCAATTG 2169
QY 699 GGTAAAGTTCATGATACCTTTATATGCGGGCTTCGCCGATCTCATGGGGTATATTCCGCTC 758

```

Db 2170 GGTAAAGTTCATGATACCCCTCACGTGCGGCTTCGCGGATCTCATGGGTATACATTCGCTC 2229
Qy 759 GTGGGCGCCCTTAGGGGCGCTGCCAGGCGCTTGGCAACAGGTGTCCGGGTCTGGAG 818
Db 2230 GTGGGCGCCCTTAGGGGCGCTGCCAGGCGCTTGGCAACAGGTGTCCGGGTCTGGAG 2289
Qy 819 GAGCGGTGAATATGACAAAGGGAATTCGCGGTTGCTCTTTCTATCTTCCTCTTG 878
Db 2290 GAGCGGTGAATATGACAAAGGGAATTCGCGGTTGCTCTTTCTATCTTCCTCTTG 2349
Qy 879 GCTCTGCTGCTGTTGACATCCAGCTTCGCTTATGAAGTGCACAAGTGTCCGG 938
Db 2350 GCTTGTGCTGCTGTTGACATCCAGCTTCGCTTATGAAGTGCACAAGTGTCCGG 2409
Qy 939 ATATACCATGTACGAACAGCTGTCTCAACTCAAGCAATGTGTATGAGGAGCGGAGCGTG 998
Db 2410 GTGTACCATGTACGAACAGCTGTCTCAAGCAATGTGTATGAGGAGCGGAGCATG 2469
Qy 999 ATCATGCATACTCCGGGTGCTGCTGTTTTCAGAGGGTAAACAGCTCCCGTTGCTG 1058
Db 2470 ATCATGCATACTCCGGGTGCTGCTGTTTTCAGAGGGTAAACAGCTCCCGTTGCTG 2529
Qy 1059 GTAGCGTCACTCCAGCTCGCGGCCAGGAATGCCAGCGTCCCACTACGACAAATACGA 1118
Db 2530 GTAGCGTCACTCCAGCTCGCGGCCAGGAATGCCAGCGTCCCACTACGACAAATACGA 2589
Qy 1119 CGCACGCTCGACTTGTCTGTTGGAGCGCTTCTGCTCGCTATGTATGAGTGGGGGAT 1178
Db 2590 CGCATGTGATTTGCTGTTGGAGCGCTTCTGCTCGCTATGTATGAGTGGGGAT 2649
Qy 1179 CTCTGCGGATCTATTTCTGCTCCAGCTGCTCACTTCTCGCTCGCGGCGCATGAG 1238
Db 2650 CTCTGCGGATCTGTTTCTGCTCGCGCATGTTCTACCTTCTCGCTCGCGGCGCATGAG 2709
Qy 1239 ACAGTGAGGACTGCACTGCTCAATCTATCCCGGCAATGTATCAGGTCAAGCATGGCT 1298
Db 2710 ACAGTGAGGACTGCAATGTCTCAATATATCCCGGCCAGCTGACAGGTCAAGCATGGCT 2769
Qy 1299 TGGGATATGATGATGAACTGTGCTACCTACACAGCCCTTAGTGGTGTGCGAGTGTCTCGG 1358
Db 2770 TGGGATATGATGATGAACTGTGCTACCTACACAGCCCTTAGTGGTGTGCGAGTGTCTCGG 2829
Qy 1359 ATCCACAAAGCTGTCTGAGCATGTGCGGGGGGCCCACTGGGGAGTCTTGGCGGGCTT 1418
Db 2830 ATCCACAAAGCTGTCTGAGTATGTTGGCGGGGCCCAATGGGGAGTCTTAGCGGGCTT 2889
Qy 1419 GCCTACTATTCCATGTTAGGGAATGTGGCTAAGGTTCTGATTTGATGCTACTCTTTGCC 1478
Db 2890 GCCTACTATTCCATGTTAGGGAATGTGGCTAAGGTTCTGATTTGATGCTACTCTTTGCC 2949
Qy 1479 GCGCTTGAGGGGAGACCCACAGCAGGGGAGGTGGCGGCCACACACCTTCGGGTTG 1538
Db 2950 GCGCTTGAGGGGGAACCTATGTGACAGGGGGAGCATGGGCAAAAACACCTTCGGGAT 3009
Qy 1539 ACGTCCCTTTCTCATCTGGGCGCTCAGAAAATCCAGCTTGTGAATACCAACGGCAGC 1598
Db 3010 ACGTCCCTTTCTCATCTGGGCGCTCAGAAAATCCAGCTTGTGAATACCAACGGCAGC 3069
Qy 1599 TGGCACATCAACAGGACTGCCCTAAATTCGAATGACTCCCTCCAACTGGGTTCTTTGCC 1658
Db 3070 TGGCACATCAACAGGACTGCCCTAAATTCGAATGACTCCCTCCAACTGGGTTCTTTGCC 3129
Qy 1659 GGGCTGTTTACGACACAAAGTTCAACTGCTCGGGTGGCGGAGCGCATGGCCAGTGC 1718
Db 3130 GGGCTGTTTACGACACAAAGTTCAACTGCTCGGGTGGCGGAGCGCATGGCCAGTGC 3189
Qy 1719 GCGCCATTTGACTGTTGCCAGGGTGGGGCCCATCACTACTAAGCTTAACAGC 1778
Db 3190 AGCCCCATGAGCGGTTGCTCAGGGTGGGGGCCCATCACTTAACATGAGTCAACAGC 3249
Qy 1779 TCGGATCAGAGGCTTATTTGCTGGCATTAACGCGCTCGACCGGTGGTGTGTTACCCGCG 1838

Db 3250 TCGGACCAAGAGCCCTTATTTGGCTAGGCACACCCCGCGCTGCGGTATCGTACCCGCG 3309
Qy 1839 TCGAGGTGTGTGCTCAGTGTATTTTCACCCCAAGCCCTGTGTGTGCTGGGACACC 1898
Db 3310 GCGCAGGTGTGTGCTCAGTGTATTTTCACCCCAAGCCCTGTGTGTGCTGGGACACC 3369
Qy 1899 GATCGTTCCGTTGCTCCTACGTATAGCTGGGGGAGAAATGAGACAGCTGATGCTCTTC 1958
Db 3370 GACCGGTTCCGCTGCTCCTACGTACAGTTGGGGGAGAAATGAGACAGCTGCTGCTCTT 3429
Qy 1959 AACAAACACGCTCCGCAACAGGCAACTGGTTGCTGTATCATGGAATGATGCTGGG 2018
Db 3430 AACAAACACGCTCCGCAACAGGCAACTGGTTGCTGTATCATGGAATGATGCTGGG 3489
Qy 2019 TTCACTAAGACGTGCGGAGTCCCCCGTGTAAACATCGGGGGTGGTAAACGACCTTG 2078
Db 3490 TTCACTAAGACGTGCGGAGTCCCCCGTGTAAACATCGGGGGATGGCAATAAACCCTTG 3549
Qy 2079 ATCTGCCCCACGAGCTGCTTCCGGAAGCACCCGAGGCTACTTACACAAATGTGGCTCG 2138
Db 3550 ACCTGCCCCACGAGCTGCTTCCGGAAGCACCCGAGGCACTTACACCAAGTGTGGTTCG 3609
Qy 2139 GGGCCCTTGTGACACCTAGTGCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCC 2198
Db 3610 GGGCCCTTGTGACACCTAGTGCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCC 3669
Qy 2199 TGCACTCTCAATTTTTCATCTTAAAGTTAGATGTATGTGGGGGCTGGAGCACAGG 2258
Db 3670 TGCACTCTCAATTTTTCATCTTAAAGTTAGATGTATGTGGGGGCTGGAGCACAGG 3729
Qy 2259 CTCAATGCGCATGCAATTTGGACTCGAGGAGAGCGGTGTAACTTTGGAGGACAGGATAGG 2318
Db 3730 CTCAATGCGCATGCAATTTGGACTCGAGGAGAGCGGTGTAACTTTGGAGGACAGGACAGA 3789
Qy 2319 TCAGAACTCAAGCCGCTGCTGTCTTCAACAGAGTGGCAGATACTGCCCTGTGCTTTC 2378
Db 3790 TCAGAGCTTACCCGCTGCTGTCTTCAACAGAGTGGCAGGTATTTGCCCTGTCTTCTC 3849
Qy 2379 ACCACCTACCGGCTTTATCCACTGGTTGATCCATCTCCATCAGAAACATCGTGGAGCTG 2438
Db 3850 ACCACCTACCGGCTTGTCTCCACTGGTTGATCCATCTCCATCAGAAACGCTCGTGGACGTA 3909
Qy 2439 CAATACCTGTACGCTGTAGGCTCAGCGTGTGCTCTCTTTGCAATCAATGGAGTACATC 2498
Db 3910 CAATACCTGTACGCTGTAGGCTCAGCGTGTGCTCTCTTTGCAATCAATGGAGTATGTC 3969
Qy 2499 CTGTTGCTTTTCTCTCTGCGAGACGCGCGGTGTGCTGCTGCTTGTGATGATGCTG 2558
Db 3970 CTGTTGCTTTTCTCTCTGCGAGACGCGCGGTGTGCTGCTGCTTGTGATGATGCTG 4029
Qy 2559 CTGATAGCCAGGCTGAGCGGCTTAGAGAACTTGGTGGTGTCTCAATGCGGCGCTCGTG 2618
Db 4030 CTGATAGCTCAAGCTGAGGCGGCTTAGAGAACTGCTGGTGTCTCAACGCGGCACTCGGTG 4089
Qy 2619 GCGGAGGCGCATGATGTTCTCTCTTCTTGTGTTCTTCTGCGCGGCTGTGATCAATTAAG 2678
Db 4090 GCGGAGGCGCATGATGTTCTCTCTTCTTGTGTTCTTCTGCGCGGCTGTGATCAATTAAG 4149
Qy 2679 GCGAGGCTGCTCTCTGCGGCGGCTATGCTTTTTATGGCGTATGGCGGCTGCTCTGCTC 2738
Db 4150 GCGAGGCTGCTCTCTGCGGCGGCTATGCTTTTTATGGCGTATGGCGGCTGCTCTGCTC 4209
Qy 2739 CTACTGCGGTTTACCAACAGCTTACGCTTGGACCGGGAGATGGGTGCAATGCGGGG 2798
Db 4210 CTGCTGGGCTTACCAACAGGATACGCGATGAGCGGAGATGGCAGCATCGTGCGGA 4269
Qy 2799 GGTGCGGTTCTTGTAGGTCTGGTATTTCTGACCTTGTCAACATCAAAAGTGTCTC 2858
Db 4270 GGTGCGGTTTCTGTTAGGTCTGATCTCTTGAACCTTGTCAACCGCACTAAGCTGTCTC 4329
Qy 2859 ACTAGGCTCATATGCTGTTACATCTTTATCAGAGCGGAGGCGCATGCAAGCTG 2918
Db 4330 GCTAGGCTCATATGCTGTTACATCTTTATCAGAGGCGGAGGACACTTTCAGCTG 4389

QY 2919 TGGGTCCCCCTCAACGTTCCGGGAGGCGCGATGCCATCATCCTCCTCAGCTGTGCG 2978
Db 4390 TGGATCCCCCTCAACGTTCCGGGGGCGCGGATGCCATCCTCCTCAGCTGCGCG 4449
QY 2979 GTTCATCCAGAGTTAAATTTTGGACATCAACAACTCTCTGCTGCCATACTCGGCCGCTC 3038
Db 4450 ATCCACCCAGAGCTAATCTTTTACCATCAACAAATCTTGTCTGCCATACTCGGTCCACTC 4509
QY 3039 ATGGTCTCCAGGCTGGCATTAACGAGATGCGGTACTTCTGTCGGCGCTCAAGGGCTCAT 3098
Db 4510 ATGGTCTCCAGGCTGGTAAACCAAGTGGCGTACTTCTGTCGGCGCACACGGGCTCAT 4569
QY 3099 CGTGATGCATGTTAGTGGAAAGTCGCGGGGTCATTATGTCCAAATGCTCTTCATG 3158
Db 4570 CGTGATGCATGCTCGTGGGAGGTGCTGGGGGTCAATATGTCCAAATGCTCTCATG 4629
QY 3159 AAGCTGGGCGCTGACAGGTACGTACGTTTATAACCATCTTACCCCACTCGGGACTGG 3218
Db 4630 AAGTTGGCGCACTGACAGGTACGTACGTTTATGACCATCTCACCCCACTCGGGACTGG 4689
QY 3219 GCCACGGGGCTACGAGACCTTGGGTGGGGGTAGAGCCGTCGTCTTCTCCGCCATG 3278
Db 4690 GCCACGGGGCTACGAGACCTTGGGTGGGGGTAGAGCCGTCGTCTTCTCTGATATG 4749
QY 3279 GAGACCAAGTTCATCACCTGGGAGCAGACCGCTGCGTGGGGACATCATCTTGGGT 3338
Db 4750 GAGACCAAGTTATCACCTGGGGGAGACACCGCGGGGTGTGGGACATCATCTTGGGC 4809
QY 3339 CTACCCGCTCTCCGCCGGAAGGGAGAGATATTTTGGGACCGGCTGATAGTCTCGAA 3398
Db 4810 CTGCCGCTCTCCGCCGAGGGGAGGAGATACATCTGGGACCGGACAGACCTTGAA 4869
QY 3399 GGGCAAGGTGGCGACTCTCTGGGCCATPCACGGGCTACTCCCAACAAACGGGGGGCTA 3458
Db 4870 GGGCAAGGTGGCGACTCTCTGGGCCATPCACGGGCTACTCCCAACAGACGGGGGCTA 4929
QY 3459 CTTGGTGCATCATCAGTACCTCAAGCGGGGACAGAACAGGTGMAAGGGAGGTT 3518
Db 4930 CTTGGTGCATCATCAGTACCTCAAGCGGGGACAGAACAGGTGMAAGGGAGGTC 4989
QY 3519 CAAAGTGTCTACCGCAACACATCTTCTGGGGACCTGTCATCAACGGCGGTGTGCTGG 3578
Db 4990 CAAAGTGTCTCACCGCAACACATCTTCTGGGGACCTGTCATCAATGGCGTGTGTTGG 5049
QY 3579 ACTGTCTACATGGCGCTGGCTCGAAGACCTTAGCGGTCCAAAGTGCATACCCAA 3638
Db 5050 ACTGTCTATCATGGTGC CGGCTCAAAGACCTTTCGGGCCAAAGGGGCCAATCAACCAA 5109
QY 3639 ATGTACCAATGTAGACCTGACCTCGTTCGGCTGGCAGGCGCCCGGGGCGCTCC 3698
Db 5110 ATGTACCAATGTAGACCAAGACCTTCGTTCGGTGGCAAGCGCCCGGGGCGGTTC 5169
QY 3699 ATGACACCATGAGCTGTGGCAGCTCGGACCTTTACTTGGTCACGAGACATGCTGATGTC 3758
Db 5170 TTGACACCATGACCTGCGGAGCTCGGACCTTTACTTGGTCACGAGGATGCCGATGTC 5229
QY 3759 ATTCCGGTGGCGGCGGAGCAGACAGAGGGAAGTCTACTTCGCCACAGGCCGCTTC 3818
Db 5230 ATTCCGGTGGCGGCGGAGCAGACAGAGGGAAGCTACTCTCCCGACCGCCGCTTC 5289
QY 3819 TACCTGAAAGGCTCTCGGGTGGTCATTTGCTTGGCTTCGGGACAGTCGTTGGGCGTC 3878
Db 5290 TACTTGAAGGCTCTTCGGGCGGTCCACTGCTCTGCCCTTCGGGGGATGCTGTGGGCATC 5349
QY 3879 TTCGGGCTGCTGTGTGACCCCGGGGCTCGGAGCGGTGACATTCATACCCGTTGAG 3938
Db 5350 TTTCCGGCTGCGGTGTGACCCCGAGGGGTTGGAAGCGGTGGAATTGTGACCCGTCGAG 5409
QY 3939 TCTATGGAATCAACATGCGGTCTCGGTCTTTCACAGACAACTCAACCCCGCGCTGTA 3998
Db 5410 TCTATGGAACCACTATGCGGTCTCCCGGTCTTCACGGAACAATCTGCTCCCTCCGCGCTA 5469

QY 3999 CCGCAGACATTTCAAAGTGGCAGCATCTGCAGCTCCTACTGCGAGCGGCAAGAGCACCAAA 4058
Db 5470 CCGCAGACATTTCAAAGTGGCCCATCTACACGCCCTACTTGGTAGCGCAGAGACCTAAG 5529
QY 4059 GTGCCGCTGGGTATGAGCCCAAGGTAACAAGGTGCTGCTGTGAACCCGCTCGTTGCC 4118
Db 5530 GTGCCGCTGGGTATGAGCCCAAGGTAACAAGGTGCTGCTGTGAACCCGCTCGTTGCC 5589
QY 4119 GCCACCTTAGGTTTGGGGGTATATGTCGAAGGCACACGGTATCGACCTTAACATCAGA 4178
Db 5590 GCCACCTTAGGTTTGGGGGTATATGTCGAAGGCACATGATGATCGACCTTAACATCAGA 5649
QY 4179 ACTGGGTAAAGGACCATTTACACGGGCGGTCCATTACGTACTCCACTATGGCAAGTTC 4238
Db 5650 ACCGGGTAGGACCATCACACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTTT 5709
QY 4239 CTTGCCACCGGTGGCTGTTCTGGGGGCGCTATGACATCATATAATGTGATGATGCCAC 4298
Db 5710 CTTGCCACCGGTGGTGTCTGCGGGGCGCTATGACATCATATAATGTGATGATGCCAC 5769
QY 4299 TCAACTGACTCGACTACATCTTGGGCATCGGCACAGTCTCGGATCGGTACCGTGCCAC 4358
Db 5770 TCAACTGACTCGACCATATCTCGGCATCGGCACAGTCTCGGACCAAGCGGACGGCT 5829
QY 4359 GGAGCGGCTCGTGTGCTCGCCACCGCTACACCTCGGGATCGGTACCGTGCCAC 4418
Db 5830 GGAGCGGACTGCTGTGCTCGCCACCGCTACCGCTCGGGATCGGTACCGTGCCACAT 5889
QY 4419 CCCAATATCGAGGAAATAGGCTGTCCAAATCGAGAGATCCCTTCTATGGCAAGGCC 4478
Db 5890 CCAACATCGAGGAGGTGGTCTCTCCAGCACTCGAGAAATCCCTTTATGGCAAGGCC 5949
QY 4479 ATCCCATTAGGGCCATCAAGGGGGGAGGATCTCATTTCTGCCATTCCAAGAGAAA 4538
Db 5950 ATCCCATTCGCGACCATCAAGGGGGGAGGACCTCATTTCTGCCATTCCAAGAGAAA 6009
QY 4539 TGTACAGCTCGCGCAAAAGCTGACAGGCTCGGATGAACGCTGAGCATATTTACCGG 4598
Db 6010 TGTGATGAGCTCGCGGAGACTATCCGGCTCGGACTCAATGTGTAGCATATTTACCGG 6069
QY 4599 GGCCTTGATGTGCTCATACCGCTATCGGAGACGCTCGTTGTGTGGCAACAGACGCT 4658
Db 6070 GGCCTTGATGTATCCGTATACCAACTAGCGGAGACGCTCATTTCTGCCATTCCAAGAGAAA 6129
QY 4659 CTAATGACGGTTTTCACGGGCGATTTTGACTCAGTGTGATCGATGCAATACATGTGTACC 4718
Db 6130 CTAATGACGGGCTTTTACGGGCGATTTTCGAGTCAGTGTGCAATACATGTGTACC 6189
QY 4719 CAGACAGTGCAGCTTCAGCTTCGGATCCCACTTCACCATTTGAGAGACGACCGTGCCCAA 4778
Db 6190 CAGACAGTGCAGCTTCAGGCTGGACCCGACCTTCACCATTTGAGAGACGACCGTGCCCAA 6249
QY 4779 GACCGGTGTGCGCTCGCAACCGCGAGGTAGAACTGCGAGGGTGGAGTGGCATCTAC 4838
Db 6250 GACCGGTGTCACTCGCAGCGGAGGAGGACTGGTAGGGGAGGATGGGCAATTTAC 6309
QY 4839 AGTGTGTGACTCCAGAGAACGGCCCTCGGCGATGTTGATCTTTCGGTCTGTGTGAG 4898
Db 6310 AGTGTGTGACTCCAGAGAACGGCCCTCGGCGATGTTGATCTTTCGGTCTGTGTGAG 6369
QY 4899 TGCTATGACGGGCTGTGTTGGTATGAGCTTCACCGCGCTGAGACCTCGGTTAGGTTG 4958
Db 6370 TGCTATGACGGGCTGTGTTGGTATGAGCTTCACCGCGCGGAGACCTCAGTTAGGTTG 6429
QY 4959 CGGGCTTACCTAAATPACACCGAGGTGCGGCTCTGCCAGGACCATCTCGAGTTCCTGGAG 5018
Db 6430 CGGGCTTACCTAAACACACAGGCGTCCCGCTCTGCCAGGACCATCTGAGTTCCTGGAG 6489
QY 5019 AGCGTCTTACAGGCTCACCCACATAGATGCCACTTCTCTGCCAGACTTAAACAGGCA 5078
Db 6490 AGCGTCTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAGGAGGCA 6549
QY 5079 GGAGACAACTTTCTTACTGTTGGCATATCAAGCTACAGTGTGCGCCAGGCTCAAGCT 5138

Db	6550	GGAGACAACTTCCCTTACCTGCTAGCATACCAAGGCTACGGTGTGGCCGACAGGGCTCAGGCT	6609
Qy	5139	CGACCTCCATCGTGGGACCAAAATGTGGAAAGTGTCTCATACGGCTGAAACCTTACACTGACAC	5198
Db	6610	CCACCTCCATCGTGGGACCAAAATGTGGAGTGTCTCATACGGCTAAAGCTTACCGTGTGAC	6669
Qy	5199	GGGCCAACCCCTGCTGTATAGGCTAGGAGCGGTCCAAAATGAGGTCACTCTCAACAC	5258
Db	6670	GGSCCAAGCCCTGCTGTATAGGCTGGAGCGGTTCAAAACGAGGTTACTACCACACAC	6729
Qy	5259	CCCATTAACATAACATCATGGCATGCTGGCTGACCTGGAGGTGCTCACTAGACCC	5318
Db	6730	CCCATAACCAATACATCATGGCATGCTGGCTGACCTGGAGGTGCTCAAGACACC	6789
Qy	5319	TGGGTGCTGTTGGCGGAGTCCCTTGACGCTTTTGGCCGCATATCTGCTGACGACAGCCAGT	5378
Db	6790	TGGGTGCTGTTAGCCGAGTCTTAGCAGCTCTGGCCGCGTATTTGCTGACACAGGCGAC	6849
Qy	5379	GTGGTCATTGTGGCGCAGGATCATCTTTGTCCGGGAAGCCAGCTGTGCTTCCGACAGGGAA	5438
Db	6850	GTGGTCATTGTGGCGAGGATCATCTTTGTCCGGAAAGCCGCGCATCATTCGCCACAGGGAA	6909
Qy	5439	GTCTCTTACAGAGTTCGATGATGGAAGAGTGTGCTCTCAAACTTCTTTACATCGAG	5498
Db	6910	GTCTCTTACCGGGAGTTCGATGATGGAAGAGTGTGCTCTCAAACTTCTTTACATCGAA	6969
Qy	5499	CAGGGAATCAGCTCGCCGAGCAATTCAGCAAAAGGCGCTCGGGTGTGTGCAACCGGCC	5558
Db	6970	CAGGGAATCAGCTCGCCGAACTTCAAACAGAGGCAATCGGGTGTGTCAACACAGCC	7029
Qy	5559	ACCAAGCAAGCGGAGGTGCTGCTCCCGTGGTGGAGTCCAAAGTGGCGAGCCCTTGAGACC	5618
Db	7030	ACCAAGCAAGCGGAGGTGCTGCTCCCGTGGTGGAAATCCAAAGTGGCGGACCCCTCGAAGCC	7089
Qy	5619	TTCTGGGCGGAAGCACTGTGGAATTTTCATCAGCGGAATACAGTACCTACGAGGCTTATCC	5678
Db	7090	TTCTGGGCGGAAGCAATGTGGAATTTTCATCAGCGGGATACAAATATTAGCAGGCTTGTCC	7149
Qy	5679	ACTCTGCTGGAACCCCGGATAGCATCATTTGATGGCATTTACAGCTTCTATCATAGC	5738
Db	7150	ACTCTGCTGGAACCCCGGATAGCATCATTTGATGGCATTTACAGGCTTATCACCAGC	7209
Qy	5739	CCGCTCACCAACCAAAACACCCCTCCTGTTTAAACATTTGGGGGATGGGTGGTGCCCAA	5798
Db	7210	CCGCTCACCAACCAACATACCCCTCCTGTTTAAACATCTGGGGGATGGGTGGCGCCCAA	7269
Qy	5799	CTCGTCTCTCCAGCGTGTGCTAGCTTTTGTGGGCGCCGCGATCGCCGAGCGGCTGTT	5858
Db	7270	CTTGTCTCTCCAGCGTGTCTTCTGCTTTTGTAGCGCCGCGCATCGCTGGAGCGGCTGTT	7329
Qy	5859	GGCAGCATAGCGCTTGGGAAGGCTCTCGTGGACATCTTGGCGGCTATGGGCAGAGGTA	5918
Db	7330	GGCAGCATAGCGCTTGGGAAGGCTCTGTGGATATTTTGGCAGGTTATGAGCAGAGGTT	7389
Qy	5919	GCCGCGCACTCGTGGCCCTTTAAGGTTCATGAGCGCGAGGTGCCCTCCACCGAGGACCTG	5978
Db	7390	GCAGCGCGCTCGTGGCCCTTTAAGGTTCATGAGCGCGAGATGCCCTCCACCGAGGACCTG	7449
Qy	5979	GTCAAATTACTCCCTGCGCATCTCTCTCTCGTGGTCCCTGCTGCTGGGGTCTGTGTGGCA	6038
Db	7450	GTAAACCTACTCCCTGCTATCTCTCTCCCTGGCGCCCTAGTGTGTGTGGGGTGTGTGGCA	7509
Qy	6039	GCAATACTGCTGCGACGTTGGGCCCGGAGAGGGGGCTGTGCAAGTGGATGAACCGGCTG	6098
Db	7510	GCGATACTGCTGCGCACTGTGGGCCCGAGGGAGGGGGCTGTGCAAGTGAACCGGCTG	7569
Qy	6099	ATAGCGTTTCTGCTTCGCGGGGTAAACCACTGCTCCCCCTAGCACTATGTGCTGAGAGCGAC	6158
Db	7570	ATAGCGTTTCTGCTTCGCGGGGTAAACCACTGCTCCCCCTAGCACTATGTGCTGAGAGCGAC	7629
Qy	6159	GCTGACGACGCTGTCACTCAGATCTCTCTTACGCTTACCATCTCAACTGCTGAAGCGG	6218

Db	7630	GCTGACGACGTTGTCACCTCAGATCCTCTCTAGCTTTACCATCACTCAGTCTCTGAAGAGG	7689
Qy	6219	CTCCACCACTGGATTAAATGAGGACTGCTCTACGCCATGCTCCGGCTCTGTGCTTAAGGAT	6278
Db	7690	CTTACCACTGGATCAACGAGGACTGCTCCAAGCCATGCTCCGGCTCTGTGCTTAAGAGAT	7749
Qy	6279	GTTTGGGATTGGATATGACACGGTGTGACTGACTTCAAGACCTGGCTCCAGTCCAAACTC	6338
Db	7750	GTTTGGGATTGGATATGACACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTC	7809
Qy	6339	CTGCGCGGTTTACCGGGAGTCCCTTTCTGTTCATGCCAAACGGGGTTACAAGGAGTCTGG	6398
Db	7810	CTGCGCGGATTGCGGGAGTCCCTTCTCTCATGTCAACGTGGGTACAAGGGAGTCTGG	7869
Qy	6399	CGGGGGGACGGCATCATGAACACCACTCCCATGGGAGCACACATCGCCGACATGTC	6458
Db	7870	CGGGGGGACGGCATCATGAACACCACTCCCATGGGAGCACACATCACCGGACATGTG	7929
Qy	6459	AAAAACGGTTCCATGAGGATCTGTAGGGCTAGAACTCGACAGCAACACGTGGCACGAAACG	6518
Db	7930	AAAAACGGTTCCATGAGGATCTGTGGGCCCTAGGACCTGTAGTAAACAGTGGCATGGAACA	7989
Qy	6519	TTCCCATCAACGCATACACCAACGGGACCTTGACACACCTCCCGCGCGCCAACTATTCC	6578
Db	7990	TTCCCATTAACGCGTACACCAACGGGCGCTCGACGCCCTCCCGCGCGCAAAATTAATCT	8049
Qy	6579	AGGGCGTATGGCGGGTGTGCTGTGAGGAGTACGTGGAGTTACCGGTGTGGGGATTTTC	6638
Db	8050	AGGGCGCTGTGGCGGGTGTGCTGTAGGAGTACGTGGAGGTTACCGCGGTGGGGGATTTTC	8109
Qy	6639	CACCTAGTGACGGGATGACCACTGACAAAGTAAAGTGCCCATGCCAGGTTCCCGGCCCCC	6698
Db	8110	CACCTAGTGACGGGATGACCACTGACAAAGTAAAGTGCCCATGCCAGGTTCCCGGCCCCC	8169
Qy	6699	GAATTTCTTCAACGAGGTGGAGTGCGGTTGCAACAGTACGCTCCGGCGTGCAAACT	6758
Db	8170	GAATTTCTTCAACGAGGTGGAGTGCGGTTGCAACAGTACGCTCCAGCGTGCAAACCC	8229
Qy	6759	CTTCTACGGAGACGTACGTTCCAGTTCGGGCTCAACCAATACTTGTTCCGGTCCGACG	6818
Db	8230	CTTCTACGGAGAGGTACATTTCTTGGTGGGCTCAATCAATACCTGGTTGGGTCCACAG	8289
Qy	6819	CTCCCATCGGAGCCGAAACCGGACGTAAACAGTGCTTACTTTCCATGTCTACCGATCCCTCC	6878
Db	8290	CTCCCATCGGAGCCGAAACCGGACGTAGCAGTGCTCACTTCCATGTCCACGACCCCTCC	8349
Qy	6879	CACATTACGACAGACGGCTAAGCGTACGCTGGCTAGAGGGTCTCCGCCCTCTTTAGCC	6938
Db	8350	CACATTACGCGGAGACGGCTAAGCGTACGCTGGCTGGCAGGGGATCTCCGCCCTCTTTGGCC	8409
Qy	6939	AGCTCATCAGCTAGCCAGTTGTCTGGCTCTTTTGAAGCGACATGCACTACCCACCAT	6998
Db	8410	AGCTCATCAGCTAGCCAGTTGTCTGGCTCTTTTGAAGCGACATGCACTACCCGTCTAT	8469
Qy	6999	GACTCCCGGACGCTGACCTCATTCGAGGCCAACTCTTGTGGCGGACAGAGATGGCGGA	7058
Db	8470	GACTCCCGGACGCTGACCTCATTCGAGGCCAACTCTTGTGGCGGACAGAGATGGCGGG	8529
Qy	7059	AACATCACTCCGCTGGAGTCAGAGATAAGTATGATTAATCTGCACTCTTTTCGAACCGCTT	7118
Db	8530	AACATCACCCCGTGGAGTCAGAAAAAAGGTAGTAAATTTTGGACTCTTTTCGAGCGGCTC	8589
Qy	7119	CACGCGAGGGGATGAGAGGGAGATATTCGTCCGGCGGAGATCTCTGCGAAATCCAGG	7178
Db	8590	CAAGCGGAGGAGTGAAGAGGAAGTATCCGTTTCCGGCGGAGATCTCTGCGGAGGTCCAGG	8649
Qy	7179	AGATTCCCTCAGCGTTGCCCATATGGSCACGCCCGGACTACAACTCTCCACTGCTAGAG	7238
Db	8650	AAATTTCCCTCGAGCGATGCCCATATGGGCAACGCCCGGATTAACAACCTCCACTGTTAGAG	8709
Qy	7239	TCTTGGAAGGACCCGGACTACGTCCCTCCAGTGGGTACACGGATGCCCATTTGCCACCTTACC	7298
Db	8710	TCTTGGAAGGACCCGGACTACGTCCCTCCAGTGGGTACACGGGTGTCTATTGCCGCTGCC	8769

QY 7299 AAGGCTCTCCAAATACCACCTCCACGGAGAAAGAGACGGTTGTCTCTGACAGAAATCCCAAT 7358
DB |||||
QY 8770 AAGGCCCTCCGATACCACTCCAGGAGGAGAGACGGTTGTCTGTGAGATCTTACC 8829
DB |||||
QY 7359 GTGTCTTCTGCTTGGCGAGCTCGCCACTAAGACCTTTGCGTAGTCCGGATCGTCGGCC 7418
DB |||||
QY 8830 GTGTCTTCTGCTTGGCGAGCTCGCCACAAAGACCTTCGGCAGCTCCGAATCGTCGGCC 8889
QY 7419 GTTGATAGCGGACCGCGACCGCCCTTCTGACCTGGCCCTCCGACGAGCGTGACAAAGGA 7478
DB |||||
QY 8890 GTTCGACAGCGCACGCGCAACGCGCTCTCTGACCAAGCCCTCCGACGACGCGACGCGGGA 8949
DB |||||
QY 7479 TCCGAGTTGAGTCGCTACTCTCCATGCCCCCTTTGAAGGGAGCGCGGGGACCCCGAT 7538
DB |||||
QY 8950 TCCGAGTTGAGTCGTACTCTCCATGCCCCCTTTGAGGGGAGCGGGGATCCCGAT 9009
QY 7539 CTCAGCGAGCGGTCTTGGTCTACCGGTGAGTGAGGAGGTAGTGAGGATGTCTCTGCTGC 7598
DB |||||
QY 9010 CTCAGCGAGCGGTCTTGGTCTACCGGTGAGTGAGGAGGTAGTGAGGACGTCTCTGCTGC 9069
QY 7599 TCAATGTCCTATACGTGAGACAGCGGCCCTTGATCAAGCCATGCGCTGCGGAGGAAAGTAAG 7658
DB |||||
QY 9070 TCGATGTCCTACACATGAGACAGCGGCCCTTGATCAAGCCATGCGCTGCGGAGGAAACCAAG 9129
QY 7659 CTGCCCCATCAACCGTTGAGCAACTCTTTGCTGCGTCAACCAACATGCTTACGCCACA 7718
DB |||||
QY 9130 CTGCCCCGTCAATGCACTGAGCAACTCTTTGCTCCGTCAACCAACTTGTCTATGCTACA 9189
QY 7719 ACATCCCGCAGCGCAAGCCTCCGGCAGAGAAAGTCACTTTTGACAGATTGCAAGTCCTG 7778
DB |||||
QY 9190 ACATCTCGCAGCGCAAGCCTCGGCCAGAGAAAGTCACTTTTGACAGACTGCGAGTCTGT 9249
QY 7779 GATGATCATTAACGGGACGTACTCAAGGAGATGAAGCGGAAAGCGGTCCACATTAAGGT 7838
DB |||||
QY 9250 GACGACCACTACCGGACGCTGTCAAGGAGATGAAGCGGAAAGCGGTCCACATTAAGGT 9309
QY 7839 AAGCTTCTATCTATAGAGAGCGCTGCAAGCTGACGCCGCCACATTCGGCCAAATCCAA 7898
DB |||||
QY 9310 AAACCTTCTATCCGTGGAGAAAGCTGTAAGCTGACGCCGCCACATTCGGGCCAGATCTAA 9369
QY 7899 TTTGGCTATGGGCAAGACGCTCCGGAACCTATCCAGCAGGCGCGTTAAACACATCCGC 7958
DB |||||
QY 9370 TTTGGCTATGGGCAAGACGCTCCGGAACCTATCCAGCAAGCGGTAAACACATCCGC 9429
QY 7959 TCCGTGTGGAGGACTTGTGGAAGACACTGAAACCAATTAACACCAATCATATGGA 8018
DB |||||
QY 9430 TCCGTGTGGAAGGACTTGTGGAAGACACTGAGACACCAATTAACACCAATCATATGGA 9489
QY 8019 AAAAGTGAGTTTCTGCTCCACACAGAGAGGAGCGGCGCAAGCCAGCTCGCCTTATC 8078
DB |||||
QY 9490 AAAAATGAGTTTCTGCTCCAAACAGAGAGAGGGGGCGCAAGCCAGCTCGCCTTATC 9549
QY 8079 GTATTCACAGACTGGGAGTTGCTGTATGCGAAGATGGCCCTTTACGACGTGCTCC 8138
DB |||||
QY 9550 GTATTCACAGATTGGGGTTGCTGTGCGAGAAATGGCCCTTTACGATGGTCTCC 9609
QY 8139 ACCCTTCTCAGCCGTGATGGGCTCCTCATACGGAATTCATATCCCCCAAGCAGCGG 8198
DB |||||
QY 9610 ACCCTTCTCAGCCGTGATGGGCTCCTCATACGGAATTCATATCCGATTCCTCTGGAAGCGG 9669
QY 8199 GTCGAGTCTCGTGAATACCTGGAAATCAAGAAATGCCCTATGGCTTCTCATATGAC 8258
DB |||||
QY 9670 GTCGAGTCTCGTGAATACCTGGAAAGCGAAATGGCCCTATGGGCTTCGCATATGAC 9729
QY 8259 ACCCGTGTTTTGACTCAACGGTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTAC 8318
DB |||||
QY 9730 ACCCGTGTTTTGACTCAACGGTCACTGAGATGACATCGGTGTTGAGGAGTCAATCTAC 9789
QY 8319 CAATGTTGACTTGGCCCCCGAGGCCAGACAGGCCAATAGGTCGTCTACAGAGCGGCTT 8378
DB |||||
QY 9790 CAATGTTGACTTGGCCCCCGAAGCCAGACAGGCCAATAGGTCGTCTACAGAGCGGCTT 9849

QY 8379 TACATCGGGGTCCCTGACTAACTCAAAGGGCAGAACTCGGTTATCGCGGTGCGC 8438
DB |||||
QY 9850 TACATCGGGGGCCCCCTGACTAAATTCATAAGGGCAGAACTCGGGCTATCGCGGTGCGC 9909
DB |||||
QY 8439 GAAAGTGGCGTGTGAGACTAGCTGGGTAAATACCTTCAATGTATCTTGAAGGCCACT 8498
DB |||||
QY 9910 GCGAGCGGTGTACTGACGACGACGCTGGGTAAATACCTTCAATGTATTTGAAGGCCGT 9969
QY 8499 GCAGCTGTGAGCTGCAGAGCTCCAGAGCTCCAGAGTGCAGTGTCTGAAACGGAGAGACACT 8558
DB |||||
QY 9970 GCGGCTGTGAGCTGCGAAGCTCCAGACTGCAAGTGTCTGATGCGGAGAGACACTT 10029
QY 8559 GTCTGTATCTGTGAAAGCGCGGAAACCCAGAGGATGCGGCGCCCTTACGAGCCTTACG 8618
DB |||||
QY 10030 GTCTGTATCTGTGAAGCGCGGGACCCCAAGAGGACGAGGCTTACGSGCCTTACG 10089
QY 8619 GAGGTATGACTAGGTATTCGCGCCCCCGGGGATTCGCGCCCAACCAAGATACGACCTG 8678
DB |||||
QY 10090 GAGGTATGACTAGATATCTGTGCCCCCTCGGGGACCCGCAACCAAGATACGACTTG 10149
QY 8679 GAGGTATGACTAGATATCTGTGCCCCCTCGGGGACCCGCAACCAAGATACGACTTG 8738
DB |||||
QY 10150 GAGGTATGACTAGATATCTGTGCCCCCTCGGGGACCCGCAACCAAGATACGACTTG 10209
QY 8739 GTATACTACCTCACCGGTGACCCCAACCCCTTGCACGCGGTGCGTGGGAGACAGCT 8798
DB |||||
QY 10210 GTGTACTATCTCACCGGTGACCCCAACCCCTTGCACGCGGTGCGTGGGAGACAGCT 10269
QY 8799 AGACACTTCAACTCAACTTTGGCTAGGCAATATCATGTATGCGCCCACTTATGG 8858
DB |||||
QY 10270 AGACACTTCAACTCAACTTTCTGCTAGGCAACATCATCATGTATGCGCCCACTTGTGG 10329
QY 8859 GCAAGGATGATCTGTAGTACTCACTTTTCTCCATCTCTTCTAGCTCAAGAGCAACTTGA 8918
DB |||||
QY 10330 GCAAGGATGATCTGTAGTACTCACTTTTCTCCATCTCTTCTAGCTCAGAAACACTTGA 10389
QY 8919 AAAGCCCTGGATGTGAGTACTAGCGGCTTGTACTCTTCTAGCCACTTGTGACCTACT 8978
DB |||||
QY 10390 AAAGCCCTGATGTGAGTACTAGCGGCTTGTACTTCTTCTAGCCACTTGTGACCTACT 10449
QY 8979 CAGATCATTTGAACGACTCCATGGTCTTAGCGCATTTACACTCCACGTTACTTCTCAGT 9038
DB |||||
QY 10450 CAGATCATTTGAACGACTCCATGGCTTAGCGCATTTTCACTTCTTCTCTCTCAGT 10509
QY 9039 GAGATCAATAGGTGGCTTCTGCTCAGGAAACTTGGGTTACACCTTGGCAACTCTG 9098
DB |||||
QY 10510 GAGATCAATAGGTGGCTTCTGCTCAGGAAACTTGGGTTACCGGCTTGGGAGTCTGG 10569
QY 9099 AGACATCGGGCCAGAGTGTCCGCTAAGCTACTGTCTCCAGGGGGGAGGGCCGCACT 9158
DB |||||
QY 10570 AGACATCGGGCCAGAGTGTCCGCTAGGCTACTGTCTCCAGGGGGGAGGGCTGCCACT 10629
QY 9159 TGTGCGAGATACCTTTTAACTGGGCAAGAAAGCAAGCTTAAACTCACTCAATTCGG 9218
DB |||||
QY 10630 TGTGCAAGTACTCTTCAACTGGGCAAGAAAGCAAGCTCAAACTCACTCAATTCGG 10689
QY 9219 GCGGCTCCAGCTGGACTTCTGCTGGTGTCTGCTGGTGTACAGGGGGGAGACATA 9278
DB |||||
QY 10690 GCTGCGTCCCAAGTGGATTTATCCAGCTGGTGTCTGCTGGTGTACAGGGGGGAGACATA 10749
QY 9279 TATCACGCTGTCTGCTGCGCCGACCCGCTGGTGTCTGCTGGTGTCTGCTTCTCT 9338
DB |||||
QY 10750 TATCACGCTGTCTGCTGCGCCGACCCGCTGGTGTCTGCTGGTGTCTCTCTCTCT 10809
QY 9339 GTAGGGGTAGGCAATTTACTCTCCCAACCGAATGAACGGGAGCTAACCTCTCAGGCC 9398
DB |||||
QY 10810 GTAGGGGTAGGCAATTTACTCTCTCCCAACCGAATGAACGGGAGCTAAACACTCAGGCC 10869
QY 9399 TTAAGCC-----ATTTCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9448
DB |||||
QY 10870 AATAGGCCATCCTGTTTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 10929
QY 9449 TTTTTCTTCTTCTTCTTTTTTTTTTCTTTTTTCTTTTTTCTTTTAAATGGTGGCTCCAT 9508

Db	10930	TTTTTTTTTCTCCTTTTTTTTTTTCCTCTCTTTTTTTTCTCTTTTCTTTTCTTTTGGTGGCTCCAT	10989
Qy	9509	CTTAGCCCTAGTCACGGCTAGCTGTGAAGGTCCTGTAGCCGCATGACTGCAGAGAGTGC	9568
Db	10990	CTTAGCCCTAGTCACGGCTAGCTGTGAAGGTCCTGTAGCCGCCTTGACTGCAGAGAGTGC	11049
Qy	9569	TGATACTGGCCCTCTCTGCAGATCATGT	9595
Db	11050	TGATACTGGCCCTCTCTGCAGATCAAGT	11076
RESULT 12			
US-09-539-601-31			
; Sequence 31, Application US/09539601C			
; Patent No. 6630343			
; GENERAL INFORMATION:			
; APPLICANT: Bartenschlager, Ralf FW			
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System			
; FILE REFERENCE: all sequences			
; CURRENT APPLICATION NUMBER: US/09/539,601C			
; CURRENT FILING DATE: 2001-08-30			
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY			
; EARLIER FILING DATE: 1999-04-03			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 31			
; LENGTH: 11076			
; TYPE: DNA			
; ORGANISM: Hepatitis C virus			
; FEATURE:			
; NAME/KEY: 5'UTR			
; LOCATION: (1)..(341)			
; OTHER INFORMATION: construct I389/Core-3'/19			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (342)..(1193)			
; OTHER INFORMATION: hepatitis C virus core - neomycin			
; OTHER INFORMATION: phosphotransferase fusion protein			
; FEATURE:			
; NAME/KEY: RBS			
; LOCATION: (1202)..(1812)			
; OTHER INFORMATION: internal ribosome entry site from			
; OTHER INFORMATION: encephalomyocarditis virus			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1813)..(10845)			
; OTHER INFORMATION: hepatitis C virus polyprotein from core to			
; OTHER INFORMATION: nonstructural protein NS5B; carries cell culture			
; OTHER INFORMATION: adaptive mutations from clone no. 19			
; FEATURE:			
; NAME/KEY: 3'UTR			
; LOCATION: (10846)..(11076)			
US-09-539-601-31			
Query Match 82.3%; Score 7894.6; DB 3; Length 11076;			
Best Local Similarity 90.8%; Pred. No. 0;			
Matches 8418; Conservative 0; Mismatches 839; Indels 10; Gaps 1;			
Qy	339	ACCATGACACGAATCCTTAACCTCAAGAAAAACCAAAAGTAACACCAACCGCGGCCCA	398
Db	1810	ACCATGGGACGAATCCTTAACCTCAAGAAAAACCAAAAGTAACACCAACCGCGGCCCA	1869
Qy	399	CAGNACGTCAAGTTCCTCCGGGCGGTGCAGATCGTTCGTGGAGTTTACCTGTTGCCCGCGC	458
Db	1870	CAGGACGTCAAGTTCCTCCGGGCGGTGCAGATCGTTCGTGGAGTTTACCTGTTGCCCGCGC	1929
Qy	459	AGGGSCCCACAGTTGGGTGTGCGCGCGACTAGGAAGCTTCCGAGCGGTGCGCAACCTCGT	518
Db	1930	AGGGSCCCACAGTTGGGTGTGCGCGCGACTAGGAAGCTTCCGAGCGGTGCGCAACCTCGT	1989
Qy	519	GGAAGGGCACCACTTATCCCAAAAGGCTCCGCGACCCGAGGCGCTGGGGCTCAGCCC	578

1659 QY GGGCTGTTTACGACACAAAGTTTCAACTCGTCCGGGTGCGGAGCGCATGSCCAGCTGC 1718
1719 Db GGGCTGTTCTACGTGCAACAAGTTTCAACTCATCTGATGCCAGAGCGATGCCAGCTGC 3189
1719 QY GCGCCCATGTAGTGGTTCGCCAGGGGTGGGCCCCCATCACTATPACTAAGCCTAACAGC 1778
1719 Db AGCCCCATCGACGCGTTCGCTCAGGGGTGGGGCCCCATCACATTACAATGAGTCAACAGC 3249
1779 QY TCGGATCAGAGCCCTTATTGCTGGCATTTACGGGCTCGACCGTGTGGTGTGATACCCGG 1838
1779 Db TCGGACCAAGAGCCCTTATTGTTGGCACTACGACACCCCGCGCTGCGGTATCGTACCCGG 3309
1839 QY TCGCAGGTGTGGTTCAGTGTATTGTTTCAACCCCAAGCCCTGTGTTGGTGGGACACAC 1898
1839 Db GCGCAGGTGTGGTTCAGTGTACTGTTTCAACCCCAAGCCCTGTGTTGGTGGGACGACC 3369
1899 QY GATCGTTCGGGTGTCCTTACGTATAGCTGGGGGAGAAATGACACAGCGTGTGTCCTC 1958
1899 Db GACCGTTTCGGGCTCCCTACGTACAGTTGGGGGAGAAATGACACGACGTGCTGCTCTT 3429
1959 QY AACAAACACGCTCCGACACAAAGGCAACTGGTTCGGCTGTACATGATGAATAGTACTGG 2018
1959 Db AACAAACACGCTCCGACGCTGCTTCGGGAGCACCCGAGGCCACTTACACCAAGTGTGTTG 3609
2019 QY GGGCCCTGTTGACACCTAGTGGTCCCTAGTAGACTACCCCATACAGGCTTTGGCACTGCC 2198
2019 Db TTCACTTAGACGTGGGAGTCCCGGTGTACATCGGGGGGTGCGTAACCGCACTTG 2078
2078 QY AACAAACACGCTCCGACGCTGTTTCAACCCCAAGCCCTGTGTTGGTGGGACGACC 3369
2078 Db AACAAACACGCTCCGACGCTGTTTCAACCCCAAGCCCTGTGTTGGTGGGACGACC 3369
2079 QY ATCTGCCCCACGCACTGCTTCCGGGAGCACCCCGAGGCTTACTTACACAAAATGTGGCTCG 2138
2079 Db ACTGCCCCACGCACTGCTTCCGGGAGCACCCCGAGGCCACTTACACCAAGTGTGTTG 3609
2139 QY GGGCCCTGTTGACACCTAGTGGTCCCTAGTAGACTACCCCATACAGGCTTTGGCACTGCC 2198
2139 Db GGGCCCTGTTGACACCTAGTGGTCCCTAGTAGACTACCCCATACAGGCTTTGGCACTGCC 3669
2199 QY TGCACCTCAATTTTCACTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 2258
2199 Db TGCACCTCAATTTTCACTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 3729
2259 QY CTCAATGCCGCAATGCAATTTGGACTCGAGGAGCGCTGTAACTTGGAGACAGGAGTAGG 2318
2259 Db CTGAAAGCCGATGCAATTTGGACTCGAGGAGCGCTGTAACTTGGAGACAGGAGTAGG 3789
2319 QY TCAGAACTCAGCCCGTGTGCTGTCTCAACAGAGTGCAGATCTGCCCTGTGCTTTC 2378
2319 Db TCAGAGCTTAGCCCGTGTGCTGTCTCAACAGAGTGCAGATCTGCCCTGTGCTTTC 3849
2379 QY ACCACCTACCGCTTTATCCACTGGTTTGAATCCATCTCCATCAGAACTCGTGACCGTG 2438
2379 Db ACCACCTACCGCTCTGTCCACTGGTTTGAATCCATCTCCATCAGAACTCGTGACCGTG 3909
2439 QY CAATACCTGTAGCGTGTAGCGTTCAGCGTTCCTCTCTTGGCAATCAAAATGGGAGTACATC 2498
2439 Db CAATACCTGTAGCGTGTAGCGTTCAGCGTTCCTCTCTTGGCAATCAAAATGGGAGTATGTC 3969
2499 QY CTGTTGCTTTTCTTCTCTGCGAGACGCGCGGTGTGCTGCTGTGTTGGATGATGCTG 2558
2499 Db CTGTTGCTTTTCTTCTCTGCGAGCGCGCGGTGTGCTGCTGTGTTGGATGATGCTG 4029
2559 QY CTGATAGCCAGGCTGAGCGCGCTTAGAGAACTTGGTGTCTCAATGGGCGCTCGTG 2618
2559 Db CTGATAGCTCAAGCTGAGCGCGCTTAGAGAACTTGGTGTCTCAATGGGCGCTCGTG 4089
2619 QY GCGGAGCGCATGCTATTCTCTCTTCTTGTGTTCTTCTGCGCCGCTGTTGATGATGCTG 2678
2619 Db GCGGAGCGCATGCTATTCTCTCTTCTTGTGTTCTTCTGCTGCTGTTGATGATGCTG 4149
2679 QY GGCAGGCTGGCTCTCGGGCGCGTATGCTTTTATGGCGTATGGCGCTGCTCTGCTC 2738
2679 Db GGCAGGCTGGCTCTCGGGCGCGTATGCTTTTATGGCGTATGGCGCTGCTCTGCTC 4209
3819 QY TACCTGAAAGGCTCCTCGGGTGGTCCATTGCTTTTGGCCCTTCGGGGGACGTCGTGGGCGTC 3878

2739 QY CTACTGGCGTTTACCAACACGAGCTTACGCCCTTGGACCGGAGATGGCTGCAATCGTGC GG 2798
2739 Db CTGCTGGCGTTTACCAACACGAGCATCGCCATGACCGGAGATGGCAGCATCGTGC GG 4269
2799 QY GGTGCGGTTCTTGTAGTCTGGTATTCTTGAACCTTGTACCATACTACAAGTGTTCCTC 2858
2799 Db GGTGCGGTTCTTGTAGTCTGGTATTCTTGAACCTTGTACCATACTACAAGTGTTCCTC 4329
2859 QY ACTAGGCTCATATGATGTTGTTACAATATCTTATACACAGAGCGGCGCACATGCAAGTG 2918
2859 Db ACTAGGCTCATATGATGTTGTTACAATATCTTATACACAGAGCGGCGCACATGCAAGTG 4389
2919 QY TGGGTCCCCCCCCCTCAACGTTTCGGGGAGCGCGATGCCATCATCTCCTCACGTGTGCG 2978
2919 Db TGGATCCCCCCCCCTCAACGTTTCGGGGAGCGCGATGCCATCATCTCCTCACGTGTGCG 4449
2979 QY GTTCATCCAGAGTTAAATTTTGAACATCAACAAATCTCTGTCTGCCATACTTGGCCGCTC 3038
2979 Db GTTCATCCAGAGTTAAATTTTGAACATCAACAAATCTCTGTCTGCCATACTTGGCCGCTC 4509
3039 QY ATGTTGCTCCAGGCTGGCATAAACGAGAGTCCGCTACTTCTGTCGCGCTCAAGGGCTCAT 3098
3039 Db ATGTTGCTCCAGGCTGGCATAAACGAGAGTCCGCTACTTCTGTCGCGCTCAAGGGCTCAT 4569
3099 QY CGTGATGATGTTAGTGCAGAAAGTCCGCGGGGTCAATTATGTCCAAATGCTCTCATG 3158
3099 Db CGTGATGATGTTAGTGCAGAAAGTCCGCGGGGTCAATTATGTCCAAATGCTCTCATG 4629
3159 QY AAGCTGGGCGGCTGACAGGTACGTATCGTTTATACCATTTTACCCCATCTGCGGAGCTGG 3218
3159 Db AAGCTGGGCGGCTGACAGGTACGTATCGTTTATACCATTTTACCCCATCTGCGGAGCTGG 4689
3219 QY GCCCAGCGGCGCTACGAGACCTTTCGGTGGCGGTAGAGCCGCTCTCTTCCGCCATG 3278
3219 Db GCCCAGCGGCGCTACGAGACCTTTCGGTGGCGGTAGAGCCGCTCTCTTCCGCCATG 4749
3279 QY GAGACCAAGGTCAATCACCCTGGGAGCAGACACCGCTGCTGTGGGAGCATCATCTTGGGT 3338
3279 Db GAGACCAAGGTCAATCACCCTGGGAGCAGACACCGCGCTGTGGGAGCATCATCTTGGGT 4809
3339 QY CTACCCGCTCTCGGCCGAGGAGGAGATATTTTGGGACCGGCTGATAGTCTCGAA 3398
3339 Db CTACCCGCTCTCGGCCGAGGAGGAGATATCTCTGGGACCGGCGAGACAGCCTTGAA 4869
3399 QY GGGCAAGGTTGGGCTCTTTCGCCCATCAGGCGCTACTCCCAACAAACCGGCGGCTGA 3458
3399 Db GGGCAAGGTTGGGCTCTTTCGCCCATCAGGCGCTACTCCCAACAAACCGGCGGCTGA 4929
3459 QY CTGTTGCTCATCATCATCTAGCTCTCAGGCGGAGCAAGAACACAGTTCGAAAGGAGGTT 3518
3459 Db CTGTTGCTCATCATCATCTAGCTCTCAGGCGGAGCAAGAACACAGTTCGAAAGGAGGTT 4989
3519 QY CAAGTGGTTTCTACCGCAACAAATCTTTCCTGGGAGCTTGATCAACGGCGTGTGCTGG 3578
3519 Db CAAGTGGTTTCTACCGCAACAAATCTTTCCTGGGAGCTTGATCAACGGCGTGTGCTGG 5049
3579 QY ACTGCTTACCATGGCGCTGGCTCGAGAGCCCTAGCGGCTTCCAAAGGTTCCAAATCACC 3638
3579 Db ACTGCTTACCATGGCGCTGGCTCGAGAGCCCTAGCGGCTTCCAAAGGTTCCAAATCACC 5109
3639 QY ATGTACACCAATGTGATGCTTGGAGCTTGGAGCGCGCTTCCCGGCGGCGGCTCC 3698
3639 Db ATGTACACCAATGTGATGCTTGGAGCTTGGAGCGCGCTTCCCGGCGGCGGCTCC 5169
3699 QY ATGACACCATGCTGTGGAGCTCGGACCTTTTACTTGGTCAAGAGCATGCTGATGTC 3758
3699 Db ATGACACCATGCTGTGGAGCTCGGACCTTTTACTTGGTCAAGAGCATGCTGATGTC 5229
3759 QY ATTCGGTGGCGGAGGCGAGCAGAGGGAAGTCTACTCTCCCGGCGGCTCC 3818
3759 Db ATTCGGTGGCGGAGGCGAGCAGAGGGAAGTCTACTCTCCCGGCGGCTCC 5289
3819 QY TACCTGAAAGGCTCCTCGGGTGGTCCATTGCTTTTGGCCCTTCGGGGGACGTCGTGGGCGTC 3878

Db	5290	TACTTGAAGGGCTCTTTCGGGCGGTCCACTGCTCTGCCCCCTCGGGGCAACGCTGTGGGCATC	5349
Qy	3879	TTCCGGGCTGCTGTGTGTCACCCCGGGGGTTCGGAAGCGGTGACATTCATACCCGTGTAG	3938
Db	5350	TTTCGGGCTGCGGTGTGCACCCGAGGGGTTCGGAAGCGGTGACATTCATACCCGTGTAG	5409
Qy	3939	TCTATGGAATACCAATGCGGTCTTCGGGTCTTCAAGAGCAATCAACCCCGGCTGTGA	3998
Db	5410	TCTATGGAATACCAATGCGGTCTTCGAGCAATCTGTCGCCCTCGGGCGGTA	5469
Qy	3999	CCGAGACATTCAGTGGCACATCTGCAGCTCTTACTGCGCAGCGGACAGACACAAA	4058
Db	5470	CCGAGACATTCAGTGGGCCATCTACAGCCCTTACTGTAAGCGGCAAGAGACTAAG	5529
Qy	4059	GTGCGGCTGCGTATGAGCCCAAGGTTACAAGGTGCTCGTCTGAACCCGCTCGTTGCC	4118
Db	5530	GTGCGGCTGCGTATGAGCCCAAGGTTACAAGGTGCTCGTCTGAACCCGCTCGTTGCC	5589
Qy	4119	GCCACCTTAGGGTTTGGGGGTATATGTCCAAGSCACACGGTATFCGACCCCTAAACATCAGA	4178
Db	5590	GCCACCTTAGGTTTCGGGGGTATATGTCTAAGSCACATGGTATCGACCCCTAAACATCAGA	5649
Qy	4179	ACTGGGTAGGACCATTAACAAGGGGCTTCATTAAGTACTCCACTATGGCAAGTTC	4238
Db	5650	ATCGGGTAAGGACCATTAACAAGGGGCTTCATTAAGTACTCCACTATGGCAAGTTC	5709
Qy	4239	CTTGCAGCGGTGCTCTGCGGGGCGCTATGACATCAATAATATGATGAGTGCAC	4298
Db	5710	CTTGCAGCGGTGCTCTGCGGGGCGCTATGACATCAATAATATGATGAGTGCAC	5769
Qy	4299	TCAACTGACTCGACTACCATCTTGGGCTCGGACAGCTCTGGACCAAGCGGACCGCT	4358
Db	5770	TCAACTGACTCGACTATCTGGGCTCGGACAGCTCTGGACCAAGCGGACCGCT	5829
Qy	4359	GGAGCGGGCTGCTGCTGCTCGCAACCGCTTACACTCGGGATCGGTTACCGTGCCACAC	4418
Db	5830	GGAGCGGGCTGCTGCTGCTCGCAACCGCTTACACTCGGGATCGGTTACCGTGCCACAT	5889
Qy	4419	CCCAATTCAGAGAAATAGGCTGCTGCAACCAATGGAGAGTCCCTTCTATGGCAAGGCC	4478
Db	5890	CCCAATTCAGAGAGGTGGCTGCTGCAACCAATGGAGAGTCCCTTCTATGGCAAGGCC	5949
Qy	4479	ATCCCCATTGAGGCATCAAGGGGGAGGATCTCATTTCTGCAATTCACCAAGAGAAA	4538
Db	5950	ATCCCCATTGAGAACATCAAGGGGGAGGACCTCATTTCTGCAATTCACCAAGAGAAA	6009
Qy	4539	TGTGACGAGCTCGCGCAAGCTGACAGGCGCTCGGACTGAACGCTGTAGCATATTACCGG	4598
Db	6010	TGTGATGAGCTCGCGCAAGCTGTCGCGCTCGGACTCAATGCTGTAGCATATTACCGG	6069
Qy	4599	GGCCTTGATGTGCGTCATACCGCTATCGGAGACGTGTTGCTGCTGGCAACAGACGCT	4658
Db	6070	GGCCTTGATGTATCGCTCATACCAATAGCGGAGACGTCATTTGCTAGCAACCGACGCT	6129
Qy	4659	CTAATGACGGGTTTACCGGCGATTTGACTCAGTGATCGACTGCAATACATGCTGACCC	4718
Db	6130	CTAATGACGGGTTTACCGGCGACTTCGACTCAGTGATCGACTGCAATACATGCTGACCC	6189
Qy	4719	CAGACAGTGCATTCAGCTTGGATCCCACTTTCACCATTTGAGACGACGACCGTGCCCAA	4778
Db	6190	CAGACAGTGCATTCAGCTTGGATCCCACTTTCACCATTTGAGACGACGACCGTGCCCAA	6249
Qy	4779	GACCGGTGTGCGTTCGCAACCGGCGAGGTAGAACTGGCAGGGGTAGAGTGGCATCTAC	4838
Db	6250	GACCGGTGTGCGTTCGCAACCGGCGAGGTAGAACTGGCAGGGGTAGAGTGGCATCTAC	6309
Qy	4839	AGGTTTGTGACTTCAGAGAAACCGGCGCTCGGCGATGTTGATTTCTCGGCTGCTGTGAG	4898
Db	6310	AGGTTTGTGACTTCAGAGAAACCGGCGCTCGGCGATGTTGATTTCTCGGCTGCTGTGAG	6369
Qy	4899	TGCTATGCGGGCTGCTGTTGATGAGCTCACGCCGCTGAGACCTCGGTTAGGTTG	4958

Db	6370	TGCTATGACGCGGGCTGCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTG	6429
Qy	4959	CGGGCTTACCTAAATACACAGAGGTTGCGGCTGTGCGAGGACCATCTGAGGTTCTGGGAG	5018
Db	6430	CGGGCTTACCTAAATACACAGAGGTTGCGGCTGTGCGAGGACCATCTGAGGTTCTGGGAG	6489
Qy	5019	AGGCTCTTACAGGCGCTCACCCACATAGATGCCACTTCTGCTCCAGACTAAACAGGCA	5078
Db	6490	AGGCTCTTACAGGCGCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGCA	6549
Qy	5079	GGAGACAACTTTCTTACCTTGGTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCT	5138
Db	6550	GGAGACAACTTTCTTACCTTGGTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCT	6609
Qy	5139	CCACTCTCATCTGTTGGGAACAAATGTGGAGTGTCTCATAGGCTGAACCTTACACTGCAC	5198
Db	6610	CCACTCTCATCTGTTGGGAACAAATGTGGAGTGTCTCATAGGCTGAACCTTACACTGCAC	6669
Qy	5199	GGGCCAACACCCCTGCTGTATAGGCTAGGAGCGCTCCAAATGAGGTATCTCTCACACAC	5258
Db	6670	GGGCCAACACCCCTGCTGTATAGGCTAGGAGCGCTCCAAATGAGGTATCTCTCACACAC	6729
Qy	5259	CCCATAACTAAATACATATGTCATGTCGCTGACCTCGAGGTGCTGCTACCTAGCACC	5318
Db	6730	CCCATAACTAAATACATATGTCATGTCGCTGACCTCGAGGTGCTGCTACCTAGCACC	6789
Qy	5319	TGGTGTGCTAGCGGAGTCTTTCAGCTTTTGGCCGCTATCTGCTGACAGAGGAGT	5378
Db	6790	TGGTGTGCTAGCGGAGTCTTTCAGCTTTTGGCCGCTATCTGCTGACAAAGGAGC	6849
Qy	5379	GTGTCATTTGCGGAGGATCATCTTGTCCGGGAAGCCAGCTGCTGCTCCGACAGAGAA	5438
Db	6850	GTGTCATTTGCGGAGGATCATCTTGTCCGGGAAGCCAGCTGCTGCTCCGACAGAGAA	6909
Qy	5439	GTCTCTTACAGGAGTTCGATGAGATGGAAGAGTGTGCTTCAAACTTCTTATCATCGAG	5498
Db	6910	GTCTCTTACAGGAGTTCGATGAGATGGAAGAGTGTGCTTCAAACTTCTTATCATCGAA	6969
Qy	5499	CAGGAAATGAGCTGCGCGAGCAATTCAGCAAAAGCGCTGCGGTTGTTGCAAAAGGCC	5558
Db	6970	CAGGAAATGAGCTGCGCGAGCAATTCAGCAAAAGCGCTGCGGTTGTTGCAAAAGGCC	7029
Qy	5559	ACCAAGCAAGGAGGCTGCTGCTCCGTTGGTGGAGTCCAGTGGCGAGCCCTTGAAGACC	5618
Db	7030	ACCAAGCAAGGAGGCTGCTGCTCCGTTGGTGGAGTCCAGTGGCGAGCCCTTGAAGACC	7089
Qy	5619	TTCTGGCGGAGCACATGTGGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCC	5678
Db	7090	TTCTGGCGGAGCACATGTGGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCC	7149
Qy	5679	ACTGCTGCGGAAACCCCGGATAGCATCATTTGATGGCATTTACAGCTTCTTATCATAGC	5738
Db	7150	ACTGCTGCGGAAACCCCGGATAGCATCATTTGATGGCATTTACAGCTTCTTATCATAGC	7209
Qy	5739	CCGCTCACACCCCAACACCCCTCTGTTTAAACATCTTGGGGGATGGTGCGTGCACAA	5798
Db	7210	CCGCTCACACCCCAACACCCCTCTGTTTAAACATCTTGGGGGATGGTGCGTGCACAA	7269
Qy	5799	CTGCTCTCTCCAGCGCTGCTGCTTCTGTCGGCGCGCGCATCGCGGAGCGGCTGT	5858
Db	7270	CTGCTCTCTCCAGCGCTGCTTCTGCTTCTGTCGGCGCGCGCATCGCGGAGCGGCTGT	7329
Qy	5859	GGCAGCATAGGCTTGGGAAGGTGCTGTCGAGCATCTTGGCGGGCTATGGGGCAGGGGTA	5918
Db	7330	GGCAGCATAGGCTTGGGAAGGTGCTGTCGAGCATCTTGGCGGGCTATGGGGCAGGGGTA	7389
Qy	5919	GCCGCGCATCTGTCGGCTTTTAAAGTTCATGAGCGCGAGGTGCCCTCCACCGAGACCTG	5978
Db	7390	GCCGCGCATCTGTCGGCTTTTAAAGTTCATGAGCGCGAGGTGCCCTCCACCGAGACCTG	7449
Qy	5979	GTCAAACTTACTCCCTGCCATCTCTCTGTCGGTCCCTGCTGTCGGGCTGCTGTCGCA	6038
Db	7450	GTCAAACTTACTCCCTGCCATCTCTCTGTCGGTCCCTGCTGTCGGGCTGCTGTCGCA	7509

QY 6039 GCAATACTGCGTCCGACACGTGGGCCCGGGAGAGGGGGCTGTGACGTGGATGAACCGGCTG 6098
Dd |||||
7510 GCGATAC1TGCCTCGGACGCTGGGCCACGGGAGGGGGCTGTGACGTGGATGAACCGGCTG 7569
QY 6099 ATAGGCTTCCGCTTCGGGGGTTAAACACAGCTCTCCCTACGCACTATGTGCTGAGAGCGAC 6158
Dd |||||
7570 ATAGGCTTCCGCTTCGGGGGTTAAACACAGCTCTCCCGCACGCACTATGTGCTGAGAGCGAC 7629
QY 6159 GCTGCAGACAGTGTCACTCAGATCTCTACGCTTACCATCACTCAACTGCTGAAGCGG 6218
Dd |||||
7630 GCTGCAGACAGTGTCACTCAGATCTCTACGCTTACCATCACTCAACTGCTGAAGAGG 7689
QY 6219 CTTCCACAGTGAATTAATGAGGACTGCTCTACGCCATGTCCGGGCTCGTGGCTAAAGGGAT 6278
Dd |||||
7690 CTTTACCAGTGTCAACGAGGACTGCTTCCAGCCATGCTCCGGCTCGTGGCTTAAGAGAT 7749
QY 6279 GTTTGGGATTGATATGACAGGTGTGACTGACTTCAAGACCTGGCTCAGTCCAAATCTC 6338
Dd |||||
7750 GTTTGGGATTGATATGACAGGTGTGACTGACTTCAAGACCTGGCTCAGTCCAAATCTC 7809
QY 6339 CTGCGCGGTTACCGGAGTCCCTTCTGCTCATGCCAACGCGGTGACAAAGGAGTCTGG 6398
Dd |||||
7810 CTGCGCGGATTGCGGGAGTCCCTTCTCTCATGTCAACGTGGGTACAAAGGAGTCTGG 7869
QY 6399 CGGGGGGACGGCATCATGCAAAACCACTGCCCATGCGAGCACAGATCGCCCGGACATGTC 6458
Dd |||||
7870 CGGGGGACGGCATCATGCAAAACCACTGCCCATGCGAGCACAGATCACCGGACATGTG 7929
QY 6459 AAAAAAGTTCATGAGGATCTGAGGCTTAAGAACTGACAGCAACACGTTGGCACCGAAG 6518
Dd |||||
7930 AAAAAAGTTCATGAGGATCTGAGGCTTAAGAACTGAGTAAACACGTTGGCATGGAACA 7989
QY 6519 TTCCCATCAACGCATACACGAGGACCTTTCGCTCATGCCAACGCGGTGACAAAGGAGTCTGG 6578
Dd |||||
7990 TTCCCATTAACGCGTACACGAGGACCTTTCGCTCATGCCAACGCGGTGACAAAGGAGTCTGG 8049
QY 6579 AGGGGCTATGCGGGTGGCTGCTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTTC 6638
Dd |||||
8050 AGGGGCTGTGGGGGTAGGTGCTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTTC 8109
QY 6639 CACTAGTGAACGGGATGACCACTGACAAACGTAAAGTGCCTTCCATGCGAGGTTCCGCGCCCG 6698
Dd |||||
8110 CACTAGTGAACGGGATGACCACTGACAAACGTAAAGTGCCTTCCATGCGAGGTTCCGCGCCCG 8169
QY 6699 GAATTTCTTACGGAGTGTAGGAGTGGTGGCTGACAGGTAAGTCCGCGGTGCAAACT 6758
Dd |||||
8170 GAATTTCTTACAGAAAGTGGATGGGTGGCTGACAGGTAAGTCCGCGGTGCAAACT 8229
QY 6759 CTTTACGGGAGGACGTCACGTTTCCAGGTCGGGCTCAACCAATATCTTGGTGGGTGCGAG 6818
Dd |||||
8230 CTTCTACGGGAGGAGTGCATCTTGGTGGGTGCGGCTCAATCAATACCTGGTGGGTGACAG 8289
QY 6819 CTTCCATGCGAGCCGAAACCGGACGTTAACGAGTCTTACCTGATGCTACCGATCCCTCC 6878
Dd |||||
8290 CTTCCATGCGAGCCCTGAAACCGGATGTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCC 8349
QY 6879 CACATTTACAGCAGACGCGCTAAGCGTAGGCTGGCTAGAGGTTCTCCCCCTCTTTAGCC 6938
Dd |||||
8350 CACATTTACGGCGGAGACGCTAAGCGTAGGCTGGCGAGGGATCTCCCCCTCTTTGGCC 8409
QY 6939 AGCTCATAGCTAGCCAGTGTCTGGGCTTCTTTGAAGGCGACATGCACTACCCACCAT 6998
Dd |||||
8410 AGCTCATAGCTAGCCAGTGTCTGGGCTTCTTTGAAGGCGACATGCACTACCCGCTCAT 8469
QY 6999 GACTCCCCGAGCGCTGACCTCATCGAGGCCAACTCTTGTGGCGGACAGATGGGCGGA 7058
Dd |||||
8470 GACTCCCCGAGCGCTGACCTCATCGAGGCCAACTCTTGTGGCGGACAGATGGGCGG 8529
QY 7059 AACNTCACTCGGCTGAGTGCAGAGATAGGTAGTAAATTTCTGGACTCTTTTGGAAACGCTTT 7118
Dd |||||
8530 AACATCACCCGCTGGAGTGCAGAAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCGGCTC 8589

QY 7119 CACGCGAGGGGATGAGAGGGAGATATCCGTGCGGCGGAGATCTCTCGGAAAAATCCAGG 7178
Dd |||||
8590 CAAGCGAGGAGGATGAGAGGGAGTATCCGTTCGCGGAGATCTCTGCGGAGTCCAGG 8649
QY 7179 AAGTTCCCTCAGCGTTGCCCATATGGGACGCCGGACTCAATCTCTCCACTGCTAGG 7238
Dd |||||
8650 AAAATTCCTCGAGCGATGCCCATATGGGACGCCGGATTACAACCTCTCACTGTTAGAG 8709
QY 7239 TCCTGGAAGGACCGGACTACGTCCCTCCGTGGTACACGGATGCCATTTGCCACTTACC 7298
Dd |||||
8710 TCCTGGAAGGACCGGACTACGTCCCTCAGTGGTACACGGGTGTCATTTGCCGCTGCC 8769
QY 7299 AAGGCTCTCCAAATACCACTCCACGGAGAAAGAGACGGTTGTCTCTGACAAATCCAA 7358
Dd |||||
8770 AAGGCCCTCCGATACCACTTACGGAGGAAGAGACGGTTGTCTCTGTCAGAAATCTACC 8829
QY 7359 GTGTCTTCTGCTTTGGCGGAGCTCGCCACTAAGACCTTTCGGTAGTCCGGATCGTGGCC 7418
Dd |||||
8830 GTGTCTTCTGCTTTGGCGGAGCTCGCCACAGAGACCTTTCGGCAGCTCCGAATCGTGGCC 8889
QY 7419 GTTGATAGCGCACGGGACCGCCCTTCTGACCTGGCTTCGACGACGGTGACAAAGA 7478
Dd |||||
8890 GTCGACAGCGGACCGGCAACGGCTCTCTGACAGCCCTCCGACGACGGCGACGGGGA 8949
QY 7479 TCCGAGTTGAGTCTGTCTCTCTGACGCCCCCTTGAAGGGGAGCCCGGGGACCCCGAT 7538
Dd |||||
8950 TCCGAGTTGAGTCTGTACTCTCTCATGCCCCCTTGAAGGGGAGCCCGGGGATCCCGAT 9009
QY 7539 CTCAGCGACGGTCTTGTCTTACCGTGTGAGTGAAGGCTGTGAGGAGTGTCTGTCTGCTG 7598
Dd |||||
9010 CTCAGCGACGGTCTTGTCTTACCGTAAAGCGAGGAGGCTGTGAGGAGCTGTCTGTCTG 9069
QY 7599 TCAATGTCTTACGTGACAGGGCCCTGTACGCGCATGCGCTCGGAGGAAAGTAAAG 7658
Dd |||||
9070 TCGATGTCTTACATGACAGAGCGCCCTGTATCGCCATGCGCTCGGAGGAAACAG 9129
QY 7659 CTGCCCATCAACCCGTTGAGCAACTTTTGTGCTGCTACCAACAAATGCTCTACGCCACA 7718
Dd |||||
9130 CTGCCCATCAATGCACTGAGCAACTTTTGTCTCGCTACCAACAACTTGGTCTATGCTACA 9189
QY 7719 ACATCCCGAGCGCAAGCTCCGGCAGAAAGGTACCTTTTGACAGATTTGCAAGTCTCTG 7778
Dd |||||
9190 ACATCTCAGCGCAAACTCTCGGCAAGAAAGGTACCTTTTGACAGATTTGCAAGTCTCTG 9249
QY 7779 GATGATCATTAACGGGAGTACTCAAGGAGATGAAGGCGTCCACAGTTAAAGGCT 7838
Dd |||||
9250 GACGACCACTACGGGACGCTCAAGGAGATGAAGGCGTCCACAGTTAAAGGCT 9309
QY 7839 AAGCTTCTATCTATAGAGGAGGCTTCAAGCTGACGCCCCCACTTCGGGCCAAATCCAAA 7898
Dd |||||
9310 AACTTCTATCCGTGAGGAGGCTTAAAGCTGACGCCCCCACTTCGGGCCAGATCTAAA 9369
QY 7899 TTTGGCTATGGGGCAAGGACGTCGGGAACTTATCCAGAGGGCGGTTAACACATCCGC 7958
Dd |||||
9370 TTTGGCTATGGGGCAAGGACGTCGGGAACTTATCCAGCAAGGCGGTTAACACATCCGC 9429
QY 7959 TCCGTTGCGGAGGACTCTCTGGAACACACTGAAACACCAATTTGACACCACTCATGSCA 8018
Dd |||||
9430 TCCGTTGGAAGGACTTCTCTGGAAGACTGAGACCAATTTGACACCACTCATGSCA 9489
QY 8019 AAAAGTGAAGTTTCTGCGTCCAAACAGAGAGGGAGCGCGCAAGCCAGCTCGCCCTTATC 8078
Dd |||||
9490 AAAATGAGTTTCTGCGTCCAAACAGAGAGGGGGCGCGCAAGCCAGCTCGCCCTTATC 9549
QY 8079 GTATTCCAGACCTCGGAGTTCGTGTATGCGAAGAGATGGCCCTTTACGACGTGCTCTCC 8138
Dd |||||
9550 GTATTCCAGATTTTGGGGGTTGCTGTGTCGAGAAAATGGGCCCTTTTACGATGTGCTCC 9609
QY 8139 ACCCTTCTCAGGCGGTGATGGGCTCTCATACGATTTTCAATACTTCCCCCAAGCAGCGG 8198
Dd |||||
9610 ACCCTTCTCAGGCGGTGATGGGCTCTTCAACGAGTTTCAATACTCTCTCTGACAGCGG 9669
QY 8199 GTCGAGTTCTCTGGTGAATACCTTGGAAATCAAGAAATGCCCTATGGGCTTCTCATATGAC 8258

Qy	582	TACCTTGGCCCTCTATGCAATGAGGGCTTGGGTGGCAGGATGGCTCTGTCAACC	641
Db	241	TATCTTTGGCCCTCTATGCAATGAGGGATGGGTGGCAGGATGGCTCTGTCAACC	300
Qy	642	CGGGCTCCCGGCCCTAGTTGGGGCCCAAGACCCCGCGGTAGGTCGCGTAATCTTGGGT	701
Db	301	CGGGCTCCCGGCCCTAGTTGGGGCCCAAGACCCCGCGGTAGGTCGCGTAATTTGGGT	360
Qy	702	AAGGTCATGATACCTTTACATGCGGCTTCGCCGATCTCATGGGGTACATTCGCGTCGTC	761
Db	361	AAGGTCATGATACCTTCATGCGGCTTCGCCGACCTCATGGGTACATTCGCGTCGTC	420
Qy	762	GGCGCCCCCTAGGGGGCGGTGCGAGGGCCTTGGCAACGCGGTGTCGGGTTCGGAGGAC	821
Db	421	GGCGCCCCCTAGGGGGCGGTGCGAGGGCCTTGGCAATGGTGTCCGGGTTCGGAGGAC	480
Qy	822	GGCGTGAACATGCAACAGGGAACTTCCCGGTTCCTTTCTATCTATCTTCCTTTGGCT	881
Db	481	GGCGTGAACATGCAACAGGGAACTTCCCGGTTCCTTTCTATCTATCTTCCTTTGGCT	540
Qy	882	CTGCTGTCTCTGTTTGACCATCCCACTTCGCGCTTATGAAGTGGCGCAAGTGTCCGGGATA	941
Db	541	CTGCTGTCTCTGCTGAACGTTACAGCTTCGCGCTCATGAAGTGGCGCAAGTGTCCGGGATA	600
Qy	942	TACCATGTCAAGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGAGCGGACGTGATC	1001
Db	601	TACCATGTCAAGAACGACTGCTCCAACTCAAGCAATGTGTATGAGGAGCGGACGTGATC	660
Qy	1002	ATGCATACTCCCGGGTGGGTGCCCTGTGTTCAGGAGGGTAAACAGTCTCCGTTGCTGGGTA	1061
Db	661	ATGCATACTCCCGGGTGGGTGCCCTGTGTTCAGGAGGGTAACTCTCCCGCTGCTGGGTA	720
Qy	1062	GGCGTCACTCCACGCTCGCGGCCAGGAATGCCAGCGTCCCACTACGACAATACGACGC	1121
Db	721	GGCGTCACTCCACGCTCGCGGCCAGGAATGCTACCATCCCACTACGACAATACGACAC	780
Qy	1122	CACGTCACTTGTCTCGTTGGAGCGGCTGCTTTCTGCTCCGCTATGTACGTGGGGATCTC	1181
Db	781	CACGTCACTTGTCTCGTTGGAGCGGCTGCTCTGCTCCGCTATGTACGTGGGGACCTC	840
Qy	1182	TGCGGATCTATTTTCTCGTCTCCAGCTGTTTCACTTCTCGCCTCGCGCGGCAATGAGACA	1241
Db	841	TGCGGATCTGTTTCTCGTCTCTCAGCTGTTTCACTTCTCGCCTCGCGCGGCAATGCGACA	900
Qy	1242	GTGCAGAGCTGCAACTGCTCAATCTATCCCGGCCATGTATCAGGTACCCGATGGCTTGG	1301
Db	901	TTGCAGAGCTGCAATGTTTCAATCTATCCCGGCCACCGGTGAGGTACCCGATGGCTTGG	960
Qy	1302	GATATGATGATGAATGCTCACTACCAAGCCCTAGTGGTGTGCGAGTTGCTCCGGATC	1361
Db	961	GACATGATGATGAATGGTCACTTCAACAGCCCTCGTAGTGTGCGAGTTTACTCCGGATC	1020
Qy	1362	CCACAAGCTGTCTGGACATGGTGGCGGGGCCCACTGGGGAGTCTTGGCGGGGCTTGCC	1421
Db	1021	CCACAAGCCGTTCATCGACATGGTGGCGGGGCCCACTGGGGAGTCTTGGCGGGGCTTGCC	1080
Qy	1422	TACTATTCCATGTTAGGAACTGGGCTAAGGTTCTGATGTTGGGCTACTCTTTTGGCGGC	1481
Db	1081	TACTATTCCATGTTGGGGGAACTGGGCTAAGGTTTGTATGTTGATGCTACTTTTGGCGGC	1140
Qy	1482	GTTGACGGGGAGACCCACACAGCGGGAGGGTGGCGGCCACACACTCCGGGTTCAAG	1541
Db	1141	GTTGACGGGGACACCTTCAACAGCGGGGGGCAAGCTGCCGCTTCAACAGGGGTTTCGG	1200
Qy	1542	TCCCTTTTCTCATTTGGGGGCTCTCAGAAAAATCCAGCTTGTGAATACCAAGCGAGCTGG	1601
Db	1201	GGCCTCTTTACCTTGCGCCGTCTCAGAGAAATCCAGCTTATAAACAACCAATGGCAGTTGG	1260
Qy	1602	CACATCAACAGACTGCGCTAAAATTGCATGACTCCCTCCAAATGGGTTCTTTGGCGGC	1661
Db	1261	CACATCAACAGGACTGCGCTGAACTGCAATGACTCCCTCCAGACTGGGTTCCTTTGGCGGC	1320
Qy	1662	CTGTTTTACGCACACAAGTTCAACTCTGCTCGGGTGGCGGAGCGCATGGCCAGCTGCGGC	1721

Db	1321	CTTTCTAGCGCATAGGTTCAACTCGTCCGATGCCGAGCGCATGGCCAGCTGCCGC	1380
Qy	1722	CCCATTTGACTGGTTCCGCCACGAGGGTGGGGCCCCCATCACCTATATCTAAGCCT---AACAGC	1778
Db	1381	TCCATTTGACAAGTTTCACACAGGATGGGGTCCATACATTATGCTGAGCCTACAAAGAC	1440
Qy	1779	TCGGATCAGAGGCTTTATTTGCTGGCATTAACGGCCCTCGACCGTGTGGTGTCTGTAACCGCG	1838
Db	1441	CCGGACAGAGGCTTATTGCTGGCACTACCCACCTCAACAATGTGTATCGTACCTGCG	1500
Qy	1839	TCCAGGTGTGGTCCAGTGTATTGTTTACCCCAAGCCCTGTTGTGTGGGGACCAACC	1898
Db	1501	TGCGAGGTGTGGTCCAGTGTATTGCTTACCCCAAGTCCCTGTGTCGTGGGGACAAAC	1560
Qy	1899	GATCGTTCCGGTGTCCCTACGTATAGTGGGGGAGAAATGAGACAGACGTGTGCTCTCTC	1958
Db	1561	GATCGTCTCGGCAACCTACGTACAGTGGGGGAGAACGATACTGACGTGCTGCTCTT	1620
Qy	1959	AACAACACGCTCCGCCACAAGCAACTGGTTCGGCTGTACATGGATGAATAGTACTGGG	2018
Db	1621	AACAACACGCGCGCGCAAGGCAACTGGTTCGGCTGTACATGGATGAATAGCACTGGG	1680
Qy	2019	TTCACTAAGACGTGCGGAGTCCCCGTGTAAACATCGGGGGGTGGTAAACCGACCTTG	2078
Db	1681	TTCAACAAAGACGTGCGGGGCCCCCGGTGTAAACATCGGGGGGTTCGGCAATAACCTTG	1740
Qy	2079	ATCTGCCCAACGACATGCTTTCGGAAAGCACCCCGAGGCTACTTTACACAAATGTGGCTCG	2138
Db	1741	ACCTGCCCAACGACATGCTTTCGGAAAGCACCCCGAGGCAACGTACTCAAAATGTGGCTCG	1800
Qy	2139	GGGCCCTGGTTGACACTAGGTGCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCC	2198
Db	1801	GGGCTTGGTTGACACTAGGTGCATGGTTGACTACCCATACAGGCTCTGGCACTACCCC	1860
Qy	2199	TGCACCTCAATTTTTCATCTTTTAAGTTAGATGTATGTGGGGGCGTGGAGCACAGG	2258
Db	1861	TGCACCTGTCAACTTCTCCATCTTTTAAGTTAGGATGTATGTGGGGGCGTGGAGCACAGG	1920
Qy	2259	CTCAATGCCGATCAATTTGGACTCGAGGAGGCGCTGTAACTTTGGAGGACAGGATAGG	2318
Db	1921	CTTAATGCTGCATGCAACTGGACCCGAGGAGAGCGTTTGCAACTTGGACGACAGGACAGA	1980
Qy	2319	TCAGAACTCAGCCGCTGCTGCTGTCTACAAAGAGTGGCAGATACTGCCCTGTGCTTTC	2378
Db	1981	TCGGAGCTCAGCCGCTGCTGCTCTCTACAAAGAGTGGCAGGTTCTGCCCTGCTCTTTC	2040
Qy	2379	ACCACTTACCGGTTTATCCACTGGTTTGATTCATCTCCATCAGAAACATCTGTGGACGTG	2438
Db	2041	ACCACTTACCGGCTGTGTCCACTGGCTTGATCCACCTCCATCAGAAACATCTGTGGACGTG	2100
Qy	2439	CAATACCTGTACCGTGTAGGCTCAGCGTTTGTCTCTCTTTCGAATCAAAATGGGAGTATGTC	2498
Db	2101	CAATACCTGTACCGTGTAGGCTCAGCGGTTGTCTCTCTTTCGAATCAAAATGGGAGTATGTC	2160
Qy	2499	CTGTTGCTTTCTCTCTCTCGCAGACGCGCGGCTGTGCTGCTCTGTGGATGATGCTG	2558
Db	2161	GTGTTGCTTTCTCTCTCTCGCAGACGCGCGCTGTGCTGCTGTGTGGATGATGCTG	2220
Qy	2559	CTGATAGCCAGGCTGAGCCGCTTAGAGAACTTGGTGTGCTCTCAATCGGCGTCCGCTG	2618
Db	2221	CTGATAGCCAGCCGAGGCGCCTTAGAGAACTTGGTGTGCTCTCAATCGGCGTCCGCTG	2280
Qy	2619	GCGGAGCGCATGGTATTCTCTCTGTTCTGTGTTCTTCTGCGCGGCTGGTACATTAAAG	2678
Db	2281	GCGGAGCGCATGGCATCTCTCTCTCTCTGTTCTTCTTGTGCGCGCTGGTACATCAAG	2340
Qy	2679	GGCAGGCTGCTCTCTGGGGCGGTATGCTTTTTTATGGCGGTATGGCGGCTGCTCCTGCTC	2738
Db	2341	GGCAGGCTGCTCTCTGGGGCGCATATGCTTTTCTATGGCGCATGGCGGCTGCTCCTGCTC	2400
Qy	2739	CTACTGCGTTTACCAACACAGCTTACGCTTGTGACCGGGAGATGGGTGCATCTGTGGG	2798

Db 2401 CTCTTGACATTAACCAACCAAGAGCTTAAGCCATGGAACCGGAGATGCTGCATGCTCGGA 2460
QY 2799 GGTGCGGTTCTGTAGGTTCTGTGATTTCTTGACCTTGTCAACCACTACTACAAAGTGTTC 2858
Db 2461 GCGCGGTTTTGTGGTCTGGCAATATTGACCTTGTGCCATATTACAAAGTGTTCCTC 2520
QY 2859 ACTAGGCTCATATGGTGGTTACAAATATTATACCAAGAGCCGAGGCGCACATGCAAGTG 2918
Db 2521 GCTAGGCTCTATGGTGGTTTACAATATCTTATCACAGAGCTGAGGCGCACTTGCATGTG 2580
QY 2919 TGGGTCCTCCCTCAAGCTTCCGGGAGGCGGATGCCATCATCTCTCTCAAGTGTGG 2978
Db 2581 TGGGTTCCCTCCCTCAAGCTCCGGGAGGCGGATGCCATCATCTCTCTCAAGTGTGA 2640
QY 2979 GTTCATCCAGAGTAAATTTTGTGACATCAACAACTCTCTGCTGCCATATCTCGGCCGCTC 3038
Db 2641 GTCCACCCAGAGCTAATCTTTGATATCAACAACTCTGATTTGCCATATCTCGGACGCTC 2700
QY 3039 ATGGTGTCCAGGCTGGCATAAACGAGAGTCCGCTACTTCTGTGGCGCTCAAGGCTCATTT 3098
Db 2701 ATGGTGTCTCAAGCTGGCATAACTAGGGTCCGCTACTTCTGTACGCGCTCAAGGCTCATTT 2760
QY 3099 CGTGCATGCATGTAGTGCAGAAAGTCCCGGGGCTCATTTATGTCCAAATGGTCTTCATG 3158
Db 2761 CGTGCATGCATGTAGTGCAGAAAGTCCCTGGGGTCAATTTATGTCCAAATGGGCTTCATG 2820
QY 3159 AAGCTGGGCGGCTGACAGAGTACGTACGTTTATAACCAATCTTACCCCACTCGCGGACTGG 3218
Db 2821 AGACTGGGCGGCTGACGGGCACTAGCTCTATAATCACTCACTCACCCCACTCGGGATTGG 2880
QY 3219 GCCACGCGGCGCTACGAGACCTTTGGGTTGGCGGTAGAGCCGCTGCTTCTCGGCCATG 3278
Db 2881 GCCACGCGGCGCTACGCGACCTTTGCGGTAGCAGTGGAGGCTGTCTCTCTGACATG 2940
QY 3279 GAGACCAAGCTCATCACTGGGAGCAGACACCGCTGCTGGGACATCATCTTGGET 3338
Db 2941 GAGACCAAGATCATCACTGGGAGGCGGACACCGCGGCTGTGGGGACATCATCTTGGET 3000
QY 3339 CTACCGCTCTCGGCGGAGGAGGAGATATTTTGGGACCGGCTGATGTCTCGAA 3398
Db 3001 CTACCTGTCTCGGCGGAGGAGGAGATATCTCTGGGCGCGCGATGTCTAGTA 3060
QY 3399 GGGCAAGGTTGGGACTCTTTGGCGCCATACCGGCTACTCTCCAAACAAACCGGGCGTA 3458
Db 3061 GGGCAGGTTGGGACTCTTTGGCGCCATACCGGCTACTCTCCAAACAAACCGGGCGTA 3120
QY 3459 CTTGGTTGCATCATCACTAGCCTCACAGGCGGAGCAAGACAGGTCGAAGGGAGGTT 3518
Db 3121 CTTGGTTGCATCATCACTAGGCTCTCACAGGCGGAGCAAGAACCCAGGTCGAGGGGAGGTT 3180
QY 3519 CAGTGGTTTCTACCGCAACAACTCTTTCTGGCGACCTGCATCAACGGGCTGTGCTGG 3578
Db 3181 CAGTGGTCTCCACCGCAACAACTCTTTCTGGGACCTGCTGCATCAACGGGCTATGTTGG 3240
QY 3579 ACTGTCTACCATGGGCTGGCTCGAAGACCTTAGCGGTCGAAAGTTCATCAACCCAA 3638
Db 3241 ACTGTCTACCATGGTGGCTCGAAGACTCTAGCGGCGGCAAGAGGCCCAATCGCCAG 3300
QY 3639 ATGTACACCAATGTAGACTGGAACCTCTGTGGCTGGGAGGCGGCGGCGGCGGCTCC 3698
Db 3301 ATGTACACTAATGTAGACAGGATCTCTGTGGCTGGCGGCGGCGGCGGCGGCTTCC 3360
QY 3699 ATGACACCATGAGTGTGGAGCTCGGACCTTTACTTGTGTGACGAGACATGCTGATGTC 3758
Db 3361 CTGACACCATGACCTGTGGAGCTTCGGACCTTTACTTGTGTGACGAGACATGCAATGTT 3420
QY 3759 ATTCCGTTGCGCGGCGGAGGCGACAGCAGGGGAGTCTACTCTCCCGGAGGCGGCTTCC 3818
Db 3421 ATTCCGTTGCGCGGCGGAGGCGACAAATAGAGGAGCTTGTCTCTCCCGGAGGCTTCTCC 3480
QY 3819 TACCTGAAGGCTCTCTCGGTTGGTCCATTTGCTTCCCTTGGGAGCAGTGTGGGCGTC 3878
Db 3481 TACTTGAAGGCTCTCTCGGTTGGCCCACTGTCTCTGCCCTTGGGAGCAGCTGTGGGCGTC 3540

QY 3879 TTCCGGGCTGCTGTGTGACACCCCGGGGGTCCGAAAGCGGTGGACTTTCATACCCGTTGAG 3938
Db 3541 TTCCGGGCGGCTGTATATGACACCCCGGGGGTTCGAAAGCGGTGGATTTTGTCCCCGTTGAG 3600
QY 3939 TCTATGGAACCTACCATGCGGTCTCCGGTCTTTACAGACAACTCAACCCCCCGGCTGTA 3998
Db 3601 TCCATGGAACCTACTATGCGGTCCCGGTCTTACAGACAACTCATCTCCCGGCGGTA 3660
QY 3999 CCGGAGACATTCGAAGTGGACATCTGACGCTCTTACTGGAGCGGCAAGACACCAA 4058
Db 3661 CCGCAACACTTCCAAGTGGCCATCTACACGCTCCCACTGGCAGCGGCAAGACACTAGA 3720
QY 4059 GTGCGGCTGCGTATGTCAGCCCAAGGTACAAGTCTCTGCTGAACCGCTCCGTTGCC 4118
Db 3721 GTGCGGCGGCGATATCGGCGCAAGGTACAAGTCTTGTCTGAAACCGCTCTGTTGCC 3780
QY 4119 GCCACCTTAGGGTTTGGGCGTATATGTCCAAGGCAACGGTATCGACCTTAACATCAGA 4178
Db 3781 GCTACCTTAGGTTTGGGCGGTATATGTCTAAGACACATGTTACCGACCTTAACATCAGG 3840
QY 4179 ACTGGGTAAGGACCAATTACCAAGGCGGCTCCATTAAGTACTTCCACCTATGCAAGTTC 4238
Db 3841 ACTGGGTAAGGACCAATTACCAAGGCGGCGGCTTACGTAATGCTGATGAGTGCAC 3900
QY 4239 CTTGCCGAGCGGTGTTCTGGGCGGCTATGACATCATATATGTGATGAGTGCAC 4298
Db 3901 CTTGCCGAGCGGTGTTCTCGGCGGCGCTTACGACATCAATATGCTGATGAGTGCAC 3960
QY 4299 TCAACTGACTTCACTTACCATCTTGGGCGCATCGGCACAGTCTCTGACCAAGCGGAGACGCT 4358
Db 3961 TCAACTGACTCACTTACTTCTTGGGCGATCGGCACAGTCTCTGACCAAGCGGAGCGCT 4020
QY 4359 GGAGCGGCTCGTGTGTGCGCAACGCTTACCTCGGGATCGGTACTCGTGCACAC 4418
Db 4021 GGAGCGGCTTGTGTGTGCGCAACGCTTACGCTTCCAGGATCGGTACCGTGCACAC 4080
QY 4419 CCATATTCAGGAAATAGSCCTGTCCAACAAATGGAGAGATCCCTTCTATGGCAAGAC 4478
Db 4081 CCATATTCAGGAGGTGGGCTGTCCAACTCTGGAGAGATCCCTTCTATCGGCAAGAC 4140
QY 4479 ATCCCATTCAGGCGCATCAAGGGGGGAGGACATCTCATTTTCTGCCATTCGAAGAGAA 4538
Db 4141 ATCCCATTCAGGCGCATCAAGGGGGGAGGACATCTCATTTTCTGTCACTCCAAGAGAG 4200
QY 4539 TGTGACGAGCTCGCGCAAGAGCTGACAGGCTCGGACTGAACGCTGTAGCATATTAACGG 4598
Db 4201 TGGACGAGCTTGGCGCAAGCTGTGAGGCTCGGAATCAATGTGTAGCTATTAACGG 4260
QY 4599 GGCCTTGAATGTCCGCTATACCGCTTATCGGAGACGTCGTTGTCTGGCAACAGACGCT 4658
Db 4261 GGTCTTGAATGTCCGCTATACCGACCGGAGACGTCGTTGTCTGGCAACAGACGCT 4320
QY 4659 CTAATGACGGGTTTACCGGCGATTTTGAATCTGATGATCGACTCAATATGATGTCAAC 4718
Db 4321 CTAATGACGGGCTATACCGGCTGATTTGATTCAGTGTGATGATGATGATGATGATGAT 4380
QY 4719 CAGACGTCGACTTCACTTGGATCCCACTTCACTTGAACGACGACGCTGCCCAAA 4778
Db 4381 CAGACGTCGACTTCACTTGGATCCCACTTCACTTGAACGACGACGCTGCCCAAA 4440
QY 4779 GACCGGTGTGCGCTCGCAACCGGCGAGGTATGAATCTGGCAGGGGTAGGATGGCATCTAC 4838
Db 4441 GACCGGTGTGCGCTCGCAGCGGCGGAGTGAATCTGGCAGGGGTAGGATGGCATATAC 4500
QY 4839 AGGTTTGTGACTCAGGAGGACGCGCTCGGCGATGTTGATTTCTGCTCTCTGTGTGAG 4898
Db 4501 AGGTTTGTGACTCAGGAGGACGCGCTCGGCGATGTTGATTTCTGCTCTCTGTGTGAG 4560
QY 4899 TGCTATGACGCGGCTGTGCTTGTGATGAGCTCAAGCGGCTGAGACCTCGGTTAGGTTG 4958
Db 4561 TGCTATGACGCGGCTGTGCTTGTGATGAGCTCAAGCGCTCAGGCGCTGAGACCTCGGTTAGGTTG 4620

QY 4959 CGGGCTTACCTAAATACACACAGGGTTGCCGTCCTCCAGGACCATCTCGAGTTCTGGGAG 5018
Db |||||
4621 CGGGCTTACCTAAATACACACAGGATTGCCCGTTTGCAGGACCATCTGAGTTCTGGGAG 4680
QY 5019 AGCGTCTTCAAGGCGCTCACCCACATAGATGCCCACTTCCTGTGTCACCACTAAACAGGCA 5078
Db |||||
4681 AGCGTCTTCAAGGCGCTCACCCACATAGATGCCCACTTCCTGTGTCACCAACAGGCA 4740
QY 5079 GGAGACAATTTCTTACTCGTGGCATATCAAGCTACAGTGTGCGCGCAGGGCTCAAGCT 5138
Db |||||
4741 GGAGATAACTTCTCCCTTACTCGTGGCATACCAAGCCACAGTGTGCGCGCAGGGCTCAGGCC 4800
QY 5139 CCACCTCCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTCAAAACCTACACTGCAC 5198
Db |||||
4801 CCACCTCCATCGTGGGATCAATGTGGAAGTGTCTCATACGGCTAAACCCACCGCTGCAC 4860
QY 5199 GGGCCAAACCCCTCTGTATAGGTAGGAGCCGTCCTCAAAATAGAGTCACTCTCACAC 5258
Db |||||
4861 GGGCCAAACGCCCTCTGTATAGGTAGGAGCCGTCCTCAAAATAGAGTCACTCTCACAC 4920
QY 5259 CCATATACTAATATACATGTCATGTCATGTCGGCTGACCTGGAGGTGCTCACTAGCACC 5318
Db |||||
4921 CCCATAACCAATATCATATGTCATGTCATGTCGGCGGACCTGGAAAGTCTGTCAACGACCC 4980
QY 5319 TGGGTGCTGTTAGGCGGAGTCTTCGAGCTTTGGCGGATACTGCTGACGACAGGCGAGT 5378
Db |||||
4981 TGGGTGCTGTTAGGCGGAGTCTTCGAGCTCTGGCGGATATTTGCTGACACAGGCGAGT 5040
QY 5379 GTGGTCATTGTGGGAGGATCATCTTGTCCGGGAAGCCAGCTGTCTGTTCCCGACAGGGAA 5438
Db |||||
5041 GTGGTTATCGTGGGTAGGATCATCTTGTCCGGGAGCGCGCTGTCTTCCCGATAGGGAA 5100
QY 5439 GTCCCTACACAGGAGTTCGATGAGATGGAGAGTGTGCTCTCAACATTCCTTACATCGAG 5498
Db |||||
5101 GTCCCTACACGGGAGTTCGATGAAATGGAAAGATGCGCTCGCACCTCCCTTACATCGAA 5160
QY 5499 CAGGGAATCGAGCTCGCCGAGCAATTCAGCAAAAGGGCTCGGGTTGTGCAACGGCC 5558
Db |||||
5161 CAGGGNATGCAACTCGCCGAGCAATTCAGCAGNAGGCGCTCGGGTTGTGCAACAGCC 5220
QY 5559 ACCAAGCAAGCGGAGGCTGCTGCTCCCGTGTGGAGTCCAACTGGCGAGCCCTTGAGACC 5618
Db |||||
5221 ACCAAGCGGCGGAGGCTGCGCTCCCGTGTGGAGTCCAACTGGCGAGCTTTGGAGACC 5280
QY 5619 TTCTGGCGGAAGCAACATGTGGAATTTTCATCAGCGGAAATACAGTACTAGCAGGCTTATCC 5678
Db |||||
5281 TTCTGGGCAAGCAAAAGTGGAAATTTTCATCAGCGGGATACAGTACTAGCGGGCTTATCC 5340
QY 5679 ACTCTGCTGGAAACCCCGGATAGCATCATTTGATGGCAATTTACAGCTTCTATCACTAGC 5738
Db |||||
5341 ACCCTGCTGGGAACCCCGGATAGCATCACTGATGGCAATTCACAGCTCTATCACAGC 5400
QY 5739 CGGCTCACCAACCAAAACCCCTCTGTTTAAACATCTTGGGGGATGGGTGCTGCCCAA 5798
Db |||||
5401 CGGCTCACCAACCAAGAACCCCTCTGTTTAAACATCTTGGGGGATGGGTGCTGCCCAA 5460
QY 5799 CTCGCTCTCCAGCGCTCGCTAGCTTTCGTGGGCGCGGATCGCCGAGCGGCTGTT 5858
Db |||||
5461 CTCGCTCCCGCAGCGCTCTTCGGGTTTCGTGGGCGCTGGTATCGCTGGTGGCGCTGTT 5520
QY 5859 GGCAGATAGGCTTGGGAAGTGTCTGAGCATCTTGGCGGCTATGGGCGAGGGTA 5918
Db |||||
5521 GGCAGATAGGCTTGGGAAGTGTCTGAGCATCTTGGCGGCTATGGGCGAGGGTG 5580
QY 5919 GCGCGGCACTCGTGGCCCTTTAAGGTCATGAGCGGAGGTCCTTCCACCGAGGACCTG 5978
Db |||||
5581 GCTGGCGCGCTCGTGGCTTCAAGGTCATGAGCGGAGCGGCCCTCTGCGCGAGGACCTG 5640
QY 5979 GTCAACTTACTCCCTGCCATCTCTCTCTGTTGCCCTGGTGTGGGGTGTGTGCGCA 6038
Db |||||
5641 ATCAATTTGCTCCCTGCCATCTCTCTCTGTTGCCCTGGTGTGTGTCGA 5700
QY 6039 GCAATACTGCTGCGGACGCTGGGCCCGGGAGGGGCTGTGCAGTGGATGAACCGGCTG 6098

Db 5701 GCAATACTGCTGCGCATGTGGGCCCGGGAGAGGGGCGGTGCAGTGGATGAACCGGCTG 5760
QY 6099 ATAGGCTTCGCTTCGCGGGGTAAACAGTCTCTCCCTAGGCACTATGCTGAGAGCCAC 6158
Db |||||
5761 ATAGGCTTCGCTTCGCGGGGTAAACATGTCTCCCCACGCACTATGCTGCGTGAAGCCAC 5820
QY 6159 GCTCAGCACGCTGTCACTCAGATCTCTTAGCCCTTACCATCACTCAACTGCTGAAGCGG 6218
Db |||||
5821 GCCGAGGGCGTGTCACTCAGGTCTCTCAGCCCTTACCATCACCCAGCTGCTGAAGAG 5880
QY 6219 CTCACACAGTGGATTAATAGGAGCTGTCTACGCCATCTCCGGCTCGTGGCTAAGGAT 6278
Db |||||
5881 CTCACACAGTGGATTAATAGGAGCTGTCTACGCCGTGTCCGGCTCGTGGCTGAGGAT 5940
QY 6279 GTTTGGGATGGATATGCACGGTGTGACTGACTTCAAGACTGTGCTCAGTCCAACTC 6338
Db |||||
5941 GTTTGGGATGGGTGTGCACGGTGTGAGTGACTTCAAGACTGTGCTCAGTCCAGTCC 6000
QY 6339 CTGCGCGGTTTACCGGGAGTCCCTTTCTGTCTATGCCAAACCGGTGCAAGGGAGTCTGG 6398
Db |||||
6001 CTGCGCGGTTTACCGGGGTCTCCCTTTCTCTATGCCAACTGGGTACAGGGAGTCTGG 6060
QY 6399 CGGGGGAACGGCATCATGCAACACCTGCCCATGCGGAGCACAGATCCCGGACATGTC 6458
Db |||||
6061 CGGGGGAACGGCATCATGCAACACCTGCCCATGCGGAGCACAGATCCCGGACATGTC 6120
QY 6459 AAAAAACGGTTCATAGGATCGTAGGGCTTAGAACCTGACAGCAACAGTGGCACGGAAG 6518
Db |||||
6121 AAAAAACGGTTCATAGGATCATCGGGCCGAAAACTGACAGCAACAGTGGCATGGAACA 6180
QY 6519 TTCCCATCAACGATACACCAACGGGACTTGTGACACCTCCCGGGCCCAACTATTTC 6578
Db |||||
6181 TTCCCATCAACGATACACCAACGGGCCCCCTGACGCCCCCTCCCGGGCCCAACTATTTC 6240
QY 6579 AGGGCGCTATGCGCGGTGGCTGCTGAGGAGTACGTGGAGGTTAACGCGTGTGGGGATTT 6638
Db |||||
6241 AAGGGCTGTGCGGGTGGCTGCTGAGGAGTACGTGGAGGTACGCGGGTGGGGATTT 6300
QY 6639 CACTACGTGACGGGATGACCACTGACAAACGTAAGTCCCATGCGAGTTCGCGGCCCC 6698
Db |||||
6301 CACTACGTGACGGGATAAACCAACGCAACACGTAAGTCCCATGTGAGTTCAGTCTCT 6360
QY 6699 GAATTTCTCAGGAGGTGGATGGGTGCGGTGACACAGTACGCTCCGCGGTGCAAACT 6758
Db |||||
6361 GAGTTTTTACGGAGGTGGATGGGGTGGGTGTGACAGGTACGCCCCCGGTGCAAACT 6420
QY 6759 CTTCTACGGGAGGAGCTCACGTTCCAGGTCCGGGTCAAACCAATATCTTGGTCCGGTCCG 6818
Db |||||
6421 CTTCTACGGGATGAGTGTGATTCAGGTCCGGCTCAATCAATACCTGTTGGGTACAG 6480
QY 6819 CTCCTATCGAGCCCGAAACCGGACGTAAACAGTGTACTTCCATGCTCACGATCCCTCC 6878
Db |||||
6481 CTCCTATCGAGCCCGAAACCGGACGTAGCAGTGTCTCACTTCCATGCTCACGACCCCTCC 6540
QY 6879 CACATTACAGCAGACAGCGCTAAGGTAGGCTAGGCTAGAGGTCTCCCCCTCTTTAGCC 6938
Db |||||
6541 CACATTACAGCAGACAGCGCTAAGGTAGGTTGGCCAGGGGTCTCCCCCTCTTTGGCC 6600
QY 6939 AGCTCATCAGCTAGCCAGTTGTCTGCGCTTCTTTTGAAGGGGACATGCACTACCC- 6993
Db |||||
6601 AGCTCTTACGTAGCCAGTGTCTGCGCCCTCTTGGAGGGGACATGCACTACCCATCT 6660
QY 6994 -----ACCATGACTCCCCGAGCGCTGACCTCATCTGAGGGCAACCTCTTGTGGCGGAG 7046
Db |||||
6661 TCCATAATCTTGACTCTCCGACGCTCGACCTCATTTGGGGGCAACCTCTGTGTGGGGCAG 6720
QY 7047 GAGATGGCGGAAACATCACTCGCGTGGAGTCAGAGTAAGTGAATTCCTGACTCT 7106
Db |||||
6721 GAGATGGCGGAAACATCACTCGCGTGGAGTCGGAGAACAGGTGTGTCTCTAGACTCT 6780
QY 7107 TTCCGAACCGCTTACCGCGAGGGGATGAGAGGGAGATATCCGTGCGCGGAGATCTCTG 7166

Db 6781 TTCGAGCCGCTTCGAGCGGAGGGGATGAGAAATATCCATTGGCGCGAGATCCTG 6840
Qy 7167 CGAAATCCAGGAAGTTCCCTCAGGGTTGCCATATGGSCACGCCCGGACTCAATTCCT 7226
Db 6841 CGGAGTCCAAAGATTCCCGCGGGATACCCATATGGGCACGGCGGGATTCAATCCT 6900
Qy 7227 CCACTGTAGAGTCTGGAAAGACCCGGACTAGTCCCTCGGTGGTACACGGATGCCA 7286
Db 6901 CCATTGTTAGAGTCTTGGAAAGAACCCGGACTAGTCCCTCGGTGGTACACGGGTGCCA 6960
Qy 7287 TTGCCACTACCAAGGCTCTCAATACCACTCCACGGAGAAAGAGACGGTTGTCCTG 7346
Db 6961 TTGCCACTGTCAAGGCCCTCCAAATACCACTCCACGGAGAAAGAGACGGTTGTCCTG 7020
Qy 7347 ACAGAAATCCAATGTGTTCTGCTTGGCGGAGCTGCCACTAAGACCTTCGGTAGCTCC 7406
Db 7021 ACGGACTCCACCGTGTTCTGTTTGGCGGAGCTCGCTACCAAAACCTTCGGCAGCTCC 7080
Qy 7407 GGATCGTGGCGGTTGATAGCGGCA CGCGGACCGCCCTTCTGACCTGGGCTCGGACGAC 7466
Db 7081 GAATTGTGGCGCGCGGACAGCGGCACCGCGACCGCCCTCTGACCAGACCTCCGACAAC 7140
Qy 7467 GGTGACAAAGATCCGAGGTGAGTCGTACTCTCCATGCCCTTGAAGGGAGCGG 7526
Db 7141 GCGGCAAAAGACTCCGACGCTGAGTCATGCTCTCTATGCCCCCTTTGAGGGGAGCGG 7200
Qy 7527 GGGACCCCGATCTCAGCGACGGGTCTTGGTCTACCGTGAGTAGGAGGCTAGTAGGAT 7586
Db 7201 GGGGACCCCGATCTCAGCGACGGGTCTTGGTCTACCGTGAGCGAGGAGGCTGGTGAGGC 7260
Qy 7587 GTCGTCTGCTCTCAATGTTCTATACGTGGACAGGCGCCCTGATCA CGCCATNGCGCTGCG 7646
Db 7261 GTCGTCTGCTCTCAATGTTCTACATGACATGGACAGGTGCCCTGATCAGGCCATGCGCGCG 7320
Qy 7647 GAGGAAAGTAAAGTCCCGATCAACCGGTTGAGCAACTCTTGTGCTGCTCACCACCAATG 7706
Db 7321 GAAGAAAGCAAGCTGCCCATCAACGCGTTGAGCAACTCTTGTGCTGCGCCATCAACAATG 7380
Qy 7707 GTCTACGCCCAACATCCCGCAGCGCAAGCCTCCGGCAGAGAAAGGTCACTTTGACAGA 7766
Db 7381 GTCTACGCCCAACATCCCGCAGCGCGGCTCGGCAGAGAGAGGTCACTTTGACAGA 7440
Qy 7767 TTGCAAGTCTCTGATGATCAATTACCGGCACTACTCAAGGAGATGAAGGCGAAGCGCTCC 7826
Db 7441 CTGCAAGTCTCTGATGATCAATTACCGGCACTGCTTAAGGAGATGAAGGCAAGCGCTCC 7500
Qy 7827 ACAGTTAAGGCTAAGCTTCTATCTATAGAGAGGCTGCAAGCTGACGCCCCGACATTCG 7886
Db 7501 ACAGTCAAGGCTAAACTTTCTATCCATAGAAGAGCCTGCCGCTGACGCCCCCACTTCG 7560
Qy 7887 GCCAAATCCAAATTTGGCTATGGGCAAGGAGCTCCGGAACCTATCCAGCAGGCGCGTT 7946
Db 7561 GCCAAATCCAAATTTGGCTATGGGCAAGGAGCTCCGGAACCTATCCAGCAGGCGCATC 7620
Qy 7947 AACCACATCCGCTCGGTGGGAGGACTTGTGGAAGACACTGGAACAACCAATTTGACACC 8006
Db 7621 AACCACATCCGCTCGGTGGGAGGACTTGTGGAAGACACTGGAACAACCAATTTGACACC 7680
Qy 8007 ACCATATGGCAAAAGTGAAGTTTCTGCGTCCAAACAGAGAGGGAGGCGCGAAGCCA 8066
Db 7681 ACCGTCATGGCAAAAGTGAAGTTTCTGCGTCCAAACAGAGAGGGAGGCGCGAAGCCA 7740
Qy 8067 GCTCGCCTTATCGTATTCAGACCTCGGAGTTCGTGTATGCGAAGATGGCCCTTTAC 8126
Db 7741 GCCCGCCTTATCGTATTCAGATTTGGAGTTTCGTGTATGCGAAGATGGCTCTCTAC 7800
Qy 8127 GACGTGCTTCCACCCTTCTCAGCGCGTGATGGGCTCCTCATACGGATTTCATATCTCC 8186
Db 7801 GATGTGTTCTCACCTTCTCTCAAGCGGTGATGGGCTCCTCATACGGATTCCAGTACTCT 7860
Qy 8187 CCCAAGCAGCGGTCAGTTCTGTTGAATACCTCGAAATCAAGAAATGCCCTATGGGC 8246
Db 7861 CCCGGGCAAGGCTCGAGTTCTGTTGTTAAAGGCTCGAAATCAAGAAACCCCTATGGGC 7920

Qy 8247 TTCTCATATGACACCGCTGTTTTGACTCAACGGTCACTGAGAGTGACATTCGTGTTGAG 8306
Db 7921 TTCTCATATGACACCGCTGTTTTGACTCAACGGTCACTGAGAATGACATTCGTGTTGAG 7980
Qy 8307 GAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAGGCCAGACAGGCCATAAGGTCGCTC 8366
Db 7981 GAGTCAATTTACCAATGTTGTGACTTTGGCCCCCGAAGCCAGACAGGCTATAAAATCGCTC 8040
Qy 8367 ACAGAGCGCTTTACATCGGGGTCCTCTGACTAACTAACTCAAAGGGCAGAACTCGCGGTTAT 8426
Db 8041 ACAGAGCGCTTTATATCGGGGTCCTCTGACTAAATTCAAAAGGGCAGAGCTTGTGTTAT 8100
Qy 8427 CGCGGTGCGCGCAAGTGGCGTGTGACGACTAGCTGCGGTAATACCTTCACATGTTAC 8486
Db 8101 CGCGGTGCGCGCAGCGCGGTGCTGACGACTAGCTGCGGTAATACCTTCACATGTTAC 8160
Qy 8487 TTCAAGGCGCACTGACGCTGTGAGCTGCAAAAGCTCAGGACTGCACGATGCTCGTGAAC 8546
Db 8161 TTCAAGGCGCTTCCGCCCTGTGAGCTGCAAAAGCTCAGGACTGCACGATGCTCGTGAAC 8220
Qy 8547 GGAGAGCAGCTTGTGCTTATCTGTAAAGCGCGGAAACCCAGAGGATGCGGGCGCCCTA 8606
Db 8221 GGGGACGACCTTGTGCTTATCTCGGAAAGCGCGGAAACCCAGAGGATGCGGGCAGCTA 8280
Qy 8607 CGAGCCTTACGAGGCTATGACTAGGTATTCCGCCCCCCCGGGGATCCGCCCAACCA 8666
Db 8281 CGAGTCTTCAAGGAGCTATGACTAGGTACTCCGCCCCCCCGGGGATTCGCCCAACCA 8340
Qy 8667 GAATACGCTGAGCTGATAACATCATGTTCTCCAATGTGTCAAGTGTGAGTGGCGACGATGCA 8726
Db 8341 GAATACGCTTGGAGTTGATPAACATCATGTTCTCCAATGTGTGCTGCGCGCAGATGCA 8400
Qy 8727 TCTGGCAAAAGGGTATACTACCTCACCGTGACCCCAACCCACCCCTTGGACGGGTGCG 8786
Db 8401 TCTGGCAAAAGGGTGTACTACCTCACTCGCGATCCCAACCCCACTCGCACGGGTGCG 8460
Qy 8787 TGGGAGACGCTAGACACACTCCAATCAACTCTTGGTAGGCAATATCATCATGTATGCG 8846
Db 8461 TGGGAAACAGCTAGACACACTCCAGTTAACTCTCTGGCTAGGCAACATTATCATGTATGCG 8520
Qy 8847 CCACCCCTATGGCAGAGGATGATTCTGATGACTCACTTTTCTCCATCTCTTAGCTCAA 8906
Db 8521 CCACCCCTATGGCAGAGGATGATTCTGATGACCAATTTCTTCCATCTCTTAGCTCAG 8580
Qy 8907 GAGCAACTTCAAAAAGCCCTGGATTGCAAAATCTACGGGCTTGTACTCAATTTAGCCA 8966
Db 8581 GAGCAACTTGA AAAAGCCCTGGATTGCAAAATCTACGGGCTTGTACTTCCATTGAGCCA 8640
Qy 8967 CTTGACCTACTCAGATCATTTGAACGACTCGATGGTCTTAGCGCATTTACACTCCAGT 9026
Db 8641 CTTGACCTACTCAGATCATTTGAAGGACTCCATGGTCTTAGCGCATTTTCACTCCATAGT 8700
Qy 9027 TACTCTCAGGTGAGATCAATAGGTGCTTCAATGCTCAGGAACTTTGGGGTACCACCC 9086
Db 8701 TACTCTCAGGTGAGATCAATAGGTGCTTCAATGCTCAGGAACTTTGGGGTACCACCC 8760
Qy 9087 TTGCAACCTTGGAGACATCGGGCCAGAAAGTGTCCCGCTTAAGCTACTGTCCAGGGGGG 9146
Db 8761 TTGGAGTCTGGAGACATCGGGCCAGGAGCGTCCCGCTTAACTACTGTCCAGGGGGG 8820
Qy 9147 AGGCGCCGCACTTGTGCGAGATACCTTTTAACTTGGGCACTTAAGGACCAAGCTTAACTC 9206
Db 8821 AGGCGCCGCACTTGTGCGCAATACCTTCTCACTGGGCACTTAAGGACCAAGCTTAACTC 8880
Qy 9207 ACTCAATCCCGCGGCTCCAGCTGCACTTGTGTGGCTGTTGCTGCTGCTGTTTACAGC 9266
Db 8881 ACTCAATCCCGGCTCCAGTCCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 8940
Qy 9267 GGGGAGACATATATCAGAGCTGCTGCTGCCGACCCCGCTGTTTCCGTTGTGCTTA 9326
Db 8941 GGGGAGACATATATCAGAGCTGCTGCTGCCGACCCCGCTGTTTTCATGCTGTGCTTA 9000

```
QY 9327 CTCCTACTTCTGTAGGGTAGGCATTTTACCTGTCCTCCCAACCGATGAACGGGAGCTAA 9386
Db 9001 CTCCTACTTCTGTAGGGTAGGCATCTACTTGTCTCCCAATCGATGAACGGGAGCTAA 9060
QY 9387 CCACTCCAGGCC-TTAAGCCATTTCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTT 9445
Db 9061 ACACCTCCAGGCCAATAGGCCATTTCCTGTTTTTTTTTTTTTTTTTTTTTTGGTTTTT 9120
QY 9446 TTTTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9500
Db 9121 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCT 9180
QY 9501 GGCTCCATCTTAGCCCTAGTCAAGCTAGCTGTGAAGGTCCGTGAGCGCATGCTGCA 9560
Db 9181 GGCTCCGCTTAGCCCTAGTCAAGCTAGCTGTGAAGGTCCGTGAGCGCATGCTGCA 9240
QY 9561 GAGAGTGTGATACCTGGCTCTCTGAGATCATGT 9595
Db 9241 GAGAGTGTGATACCTGGCTCTCTGAGATCATGT 9275

RESULT 14
US-08-324-977-13
; Sequence 13, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McIeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9030 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..9030
; OTHER INFORMATION: /note: "sequence = 333 - 9362 of
; OTHER INFORMATION: SEQ ID NO: 1"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..9030
; US-08-324-977-13

Query Match 81.4%; Score 7809.2; DB 2; Length 9030;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 8267; Conservative 0; Mismatches 763; Indels 0; Gaps 0;

QY 342 ATGAGCACGAATCCTAAACCTCAAAAGAAACCAAAACGTAACACCAACCGCCGCCACAG 401
Db 1 ATGAGCACGAATCCTAAACCTCAAAAGAAACCAAAACGTAACACCAACCGCCGCCACAG 60
QY 402 GACGTCAAGTTCCCGGGGGTGGTCAGATCGTTGGTGGAGTTTACCTGTTCCCGGGCAGG 461
Db 61 GACGTCAAGTTCCCGGGGGTGGTCAGATCGTTGGTGGAGTTTACCTGTTCCCGGGCAGG 120
QY 462 GGCCCCAGGTTGGTGGTGGCGGACTAGGAAGCTTCCGAGCGGTGCGCAACCTCGTGA 521
Db 121 GGCCCCAGGTTGGTGGTGGCGGCGCCAGGAAGCTTCCGAGCGGTGCGCAACCTCGTGA 180
QY 522 AGGCGCAACACCTATCCAAAGGCTCGCCGACCCGAGGGCGAGGGCTTGGGCTCAGCCCGGG 581
Db 181 AGGCGCAACACCTATCCCAAGGCTCGCGGCGCCGAGGGCGAGGCTTGGGCTCAGCCCGGG 240
QY 582 TACCTTTGGCCCCCTCTATGGCAATGAGGGCTCGGGGTGGGCGAGGATGGCTCTGTCAACC 641
Db 241 TACCTTTGGCCCTCTCTATGGCAATGAGGGCTTAGGGTGGGCGAGGATGGCTCTGTCAACC 300
QY 642 CGCGCTCCCGGCGCTAGTTGGGGCCCGACGACCCCGGCGCTAGTTCGCGTAACCTTGGGT 701
Db 301 CGCGCTCCCGGCGCTAGTTGGGGCCCGACGACCCCGGCGCTAGTTCGCGTAACCTTGGGT 360
QY 702 AAGGTCATCGATACCTTTACATGCGGGCTTCGCCGATCTCATGGGGTACATTCGGCTCGTC 761
Db 361 AAGGTCATCGATACCTTCACATGCGGGCTTCGCCGATCTCATGGGGTACATTCGGCTCGTC 420
QY 762 GGCGCCCCCTAGGGGGCGCTGCCAGGGCTTGGGCAACGGTGTCCGGGTTCTGGAGGAC 821
Db 421 GGCGCCCCCTAGGGGGCGCTGCCAGGGCTTGGGCAACGGTGTCCGGGTTCTGGAGGAC 480
QY 822 GCGGTGAACACTATGCAACAGGGAACCTTGCCTGGTTCCTTCTCTATCTCTCTCTCTCTCT 881
Db 481 GCGGTGAACACTATGCAACAGGGAACCTTGCCTGGTTCCTTCTCTATCTCTCTCTCTCTCT 540
QY 882 CTGCTGTCTCTTTTGACCATCCAGCTTCCGCTTATGAAGTGCAGCAACGCTGTCCGGGATA 941
Db 541 CTGCTGTCTCTGCTGACCCAGCTTCCGCTTACGAAGTGCAGCAACGCTGTCCGGGATA 600
QY 942 TACCATGTCAGAACGATGCTCCAACTCAAGACTTGTGTATGAGGCGGACGCTGATC 1001
Db 601 TATCATGTCAGAACGATGCTCCAACTCAAGACTTGTGTATGAGGCGGACGCTGATC 660
QY 1002 ATGCATCTCCCGGGTGGTGGCTGTGTTCCAGAGGGTAAACAGCTCCGCTTGTGGGTA 1061
Db 661 ATGCATCTCTGGTGGTGGCTGTGTTCCAGAGGGTAAACAGCTCCGCTTGTGGGTA 720
QY 1062 GCGCTCACTCCACGCTCGCGGCCAGGAATCCAGCGTCCCACTACGACAATACGACGC 1121
```


Db 5101 |||||CTCTACAGAGTTCGATAAAATGGAAGATGCGCTCGCACCTCCTTACATCGAGCAG 5160
QY 5502 GGAATGACGCTCGCCAGCAATTCAGCAAAAGGCGCTCGGTTGTTGCAAAAGCGCACCC 5561
Db 5161 GGAATGACGCTCGCCAGCAATTCAGCAAAAGGCGCTCGGTTGTTGCAAAAGCGCACCC 5220
QY 5562 AAGCAAGCGAGGCTGCTGCTCCGTTGAGTCCAAAGTGGCGAGCCCTTTGAGACTTTC 5621
Db 5221 AAACAAGCGAGGCTGCTGCTCCGTTGAGTCCAAAGTGGCGAGCCCTTTGAGACTTTC 5280
QY 5622 TGGCGAAGCACAATGTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTTATCCACT 5681
Db 5281 TGGCGAAGCACAATGTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTTATCCACT 5340
QY 5682 CTGCTCGGAACCCCGGATAGCATCATTTGATGSGATTTACAGCTTCTATCACTAGCCCG 5741
Db 5341 CTGCTCGGAACCCCGGATAGCATCATTTGATGSGATTTACAGCTTCTATCACTAGCCCG 5400
QY 5742 CTCACCAACCAAAACACCTCTCTGTTTAACTCTTGGGGGATGGGTGGCTGCCCAATC 5801
Db 5401 CTCACCAACCAAAAGTACCTCTCTGTTTAACTCTTGGGGGATGGGTGGCTGCCCAATC 5460
QY 5802 GCTCTCTCCAGCGCTGCGTCAGCTTTGTTGGGCGCGGCATCGCGAGCGGCTGTTGGC 5861
Db 5461 GCCCCCCCAGCGCGCTTCGGCTTTCTGTTGGGCGCGGCATCGCGGCTGTTGGC 5520
QY 5862 AGCATAGGCTTGGGAAGTGTCTGTGGAATCTTGGCGGGCTATGCGGAGGCTGAGCC 5921
Db 5521 AGCATAGGCTTGGGAAGTGTCTGTGGAATCTTGGCGGGTATGCGAGCAGGATGGCC 5580
QY 5922 GCGGCACTGTGGCTTTAAGGTATGAGCGGCGAGTGCCCTCCACCGAGGACCTGGTC 5981
Db 5581 GCGGCGCTGTGGCTTTAAGGTATGAGCGGCGAGTGCCCTCCACCGAGGACCTGGTC 5640
QY 5982 AACTTACTCTCGCATCTCTCTCTGTTGCTGCTGCGGGTGTGTGCGCAGCA 6041
Db 5641 AATCTACTCTCTCGCATCTCTCTCTGTTGCGGCCCTGTGCTGCGGGTGTGTGCGCAGCA 5700
QY 6042 ATACTGCTCGGCACGTGGGCGCGGAGAGGGGCTGTGAGTGGATGAACCGGCTGATA 6101
Db 5701 ATACTGCTCGACACGTGGGTCGGGAGAGGGGCTGTGAGTGAACCGGCTGATA 5760
QY 6102 GCGTTGCGCTTGGGGGTAAACAGTCTCCCTTACGCATATGTGCTGAGAGGAGCGCT 6161
Db 5761 GCGTTGCGCTTGGGGGTAAATCATGTTTCCCCACGCACATATGTGCTGAGAGCGAGCC 5820
QY 6162 GCAGCAGTGTCACTAGATCTCTCTAGCTTACCATCACTCACTGCTGAGCGGCTC 6221
Db 5821 GCAGCGGTGTACTCAGATCTCTCTCAGCGCTTACCATCACTCAGCTGCTGAAAAGGCTC 5880
QY 6222 CACAGTGGATTAATGAGGACTGCTTACGCCATGCTCCGSGCTCGTGGCTAAGGATGTT 6281
Db 5881 CACAGTGGATTAATGAGGACTGCTTCAACACCGTGTCCGSGCTCGTGGCTAAGGATGTT 5940
QY 6282 TGGGATTTGATATGACCGGTTTGACTGACTTTCAAGACCTGGCTCCAGTCCAAACTCCTG 6341
Db 5941 TGGGACTGGATATGCACGGTGTGACTGACTTCAAGACCTGGCTCCAGTCCAAAGCTCCTG 6000
QY 6342 CCGCGGTTTACCGGAGTCCCTTTTCTGTATGCCAACCGCGGTACAAGGAGTCTGGCGG 6401
Db 6001 CCGCAGCTTACCTGGAGTCCCTTTTCTGTTGCGCAACCGCGGTACAAGGAGTCTGGCGG 6060
QY 6402 GGGAGCGCATCATGCAAAACCACTGCGGATGCGGAGCAGAGTCCGCGGACATGTCAAA 6461
Db 6061 GGAGACGGCATCATGCAAAACCACTGCGGATGCGGAGCAGAGTCAACCGGACATGTCAAA 6120
QY 6462 AACGGTTCATGAGATCGTAGGCGCTAGAACCTTCGACACACGTGCGACGGAACGTTTC 6521
Db 6121 AACGGTTCATGAGATCGTAGGCGCTAAGACCTTCGACACACGTGCGATGGAACATTC 6180
QY 6522 CCCATCAACGCATACACCAAGGACCTTTGCAACCTCCCGCGGCGCCCACTATTTCAGG 6581
Db ||||| 7320

Db 6181 CCATCAACGCATACACACGGGCCCTGTCACACCTCTCCAGCGCCAAACTATTCTAGG 6240
QY 6582 GCGCTATGCGGGTGGCTGCTGAGGAGTACGTGGAGTTACGCGTGTGGGAGATTTCCAC 6641
Db 6241 GCGCTGTGGCGGTGGCGCTGAGGAGTACGTGGAGTCAACGCGGTGGGAGATTTCCAC 6300
QY 6642 TAGTGACGGGATGACCACTGACACGTAAAGTGCCTGACAGGTTCCGGCCCCGGA 6701
Db 6301 TAGTGACGGGATGACCACTGACACGTAAAGTGCCTGACAGGTTCCGGCTCTGAA 6360
QY 6702 TTCTTTCACGAGTGGATGGAGTGGCGTTGCAAGTACGCTCCGCGGTGCAAACTCTT 6761
Db 6361 TTCTTTCGAGGTGACGAGTGGCGTTGCAAGTACGCTCCGCGGTGCAAGGCTCTC 6420
QY 6762 CTACGGAGAGCTCACTGTTCCAGGTCGGGCTCAACAAATACATTTGTTGGTTCGAGCTC 6821
Db 6421 CTACGGAGAGGTTACATTTCCAGGTGGGCTCAACAAATACATTTGTTGGTTCACAGCTA 6480
QY 6822 CANTGCGAGCCGAAACCGGACGTAAACAGTGTCTTACCTTCCATGCTACCGATCCCTCCCAC 6881
Db 6481 CCATGCGAGCCGAAACCGGATGTAGCAGTGTCTCACTTCCATGCTACCGACCCCTCCCAC 6540
QY 6882 ATTTACGACGAGACGCTAAGCGTAGGCTAGAGGTTCTCCCGCTCTTTTAGCCAGC 6941
Db 6541 ATCAGACGAGAAACGCGTAGGTTGGCGCAGGGGTTCTCCCGCTCTTTGGCCAGC 6600
QY 6942 TCATCAGCTAGCCAGTTGTTCTGCGCTTCTTTTGAAGCGACATGCACTACCCACCATGAC 7001
Db 6601 TCTTTCAGCTAGCCAGTTGTTCTGCGCTTCTTTGAGGGACATGCACTACCCACCATGTC 6660
QY 7002 TCCCCGAGCGCTCACTCATCGAGGCCAACCTCTTGTGGCGGAGAGATGGCGGGAAC 7061
Db 6661 TCTCCGAGCGCTGACCTCATCGAGGCCAACCTCTTGTGGCGGAGAGATGGCGGGAAC 6720
QY 7062 ATCACTCGGTGAGTCAAGAAATAGGTAGTAAATTTCTGAGCTCTTTTCGAACCGCTTAC 7121
Db 6721 ATCACTCGGTGAGTCAAGAAATAGGTAGTAAATTTCTGAGCTCTTTTCGAACCGCTTTCGA 6780
QY 7122 GCGAGGGGATGAGAGGAGATACGTCGCGCGGAGATCCTGCGAAATCCAGGAAG 7181
Db 6781 GCGAGGAGGATGAGAGGAGATACGTCCTCGCGGAGATCCTGCGAAATCCAGGAAG 6840
QY 7182 TTCCCTCAGCGTTGCCATATGCGCACGCGCCGACATACAAATCTCCACCTGCTAGAGTCC 7241
Db 6841 TTCCCGCAGCATGCCCCATCTGGCGCGCCGAGTATCAACCTTCACTGTTAGAGTCC 6900
QY 7242 TGAAGGACCCGAGTACGTCCTCCGCGGTGTACACGAGTGCCTATGCCACCTTACCAAG 7301
Db 6901 TGAAGGACCCGAGTACGTCCTCCGCGGTGTGTGTCAGCGGTGCGCGTTGCCACCTATCAAG 6960
QY 7302 GCTCTCTCAATACACCTCCACGAGAAAGAGACGTTGTCTTACAGAAATCCAAATGTG 7361
Db 6961 GCGCTCTCAATACACCTCCACGAGAAAGAGACGTTGTCTTACAGAAATCCAAATGTG 7020
QY 7362 TCTTCTGCTTGGCGGAGCTGCGACCTTAAGACCTTCCGTAGCTCCGATCGTGGCGGTT 7421
Db 7021 TCTTCTGCTTGGCGGAGCTGCGACCTTCCGTAGCTCCGAGCTCCGATCATCGGCGTC 7080
QY 7422 GATAGCGGACCGGACCGGCTTCTTCTGACCTGCGCTCCGACGAGCTGACAAAGGATCC 7481
Db 7081 GACAGCGGACCGGACCGGCTTCTTCTGACGAGGCTTCCGACGAGCTGACAAAGGATCC 7140
QY 7482 GACGCTTCACTGCTTCTTCTTCTTCTTGAAGGGAGCGGGGAGCCCGATCTC 7541
Db 7141 GACGCTTCACTGCTTCTTCTTCTTCTTGAAGGGAGCGGGGAGCCCGATCTC 7200
QY 7542 AGCGACGGGTTTGGTCTACCGTGAAGTGAAGGATGTCTGTCTGTCTCA 7601
Db 7201 AGTGAAGGTTTGGTCTACCGTGAAGGATGTCTGTCTGTCTCA 7260
QY 7602 ATGTCTTATAGTGAAGGCGGCTGATCAGGCTGCGTGGGAGGAAAGTAAGCTG 7661
Db 7261 ATGTCTTATAGTGAAGGCGGCTGATCAGGCTGCGTGGGAGGAAAGTAAGCTG 7320

FILING DATE: 424
PRIOR APPLICATION DATA: US 07/769,996
APPLICATION NUMBER: 02-OCT-1991
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA: JP 2-230921
APPLICATION NUMBER: 31-AUG-1990
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA: JP 2-305605
APPLICATION NUMBER: 09-NOV-1990
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA: US 07/635,451
APPLICATION NUMBER: 28-DEC-1990
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 9030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA from genomic RNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..9030
OTHER INFORMATION: /note: "sequence = 333 - 9362 of
OTHER INFORMATION: SEQ ID NO: 1"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..9030
US-08-384-616-13

Query Match 81.4%; Score 7809.2; DB 2; Length 9030;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 8267; Conservative 0; Mismatches 763; Indels 0; Gaps 0;

Qy 342 ATGAGCAGCAATCTAAACCTCAAGAAACCAAAACGTAACACCAACCGCCGCCACAG 401
Db 1 ATGAGCAGCAATCTAAACCTCAAGAAACCAAAACGTAACACCAACCGCCGCCACAG 60

Qy 402 GACGTCAAGTTCCCGGGCGGTGGTCAGATCGTTGGTGAGTTTACCTGTTCCCGCGCAGG 461
Db 61 GACGTCAAGTTCCCGGGCGGTGGTCAGATCGTTGGTGAGTTTACCTGTTCCCGCGCAGG 120

Qy 462 GSCCCAGTTGGGTGTCGCGCGACTAGGAAGCTTCCGAGCGGTCCAAACCTCGTGA 521
Db 121 GSCCCAGTTGGGTGTCGCGCGCCAGGAAGCTTCCGAGCGGTCCAAACCTCGTGA 180

Qy 522 AGGCGACAACCTATCCCAAGGCTCGCGACCCGAGGCGAGGCTGGGGCTCAGCCCGGG 581
Db 181 AGGCGACAACCTATCCCAAGGCTCGCGCGCCGAGGCGAGGCTGGGGCTCAGCCCGGG 240

Qy 582 TACCTTGGCCCTCTATGGCAATAGAGGCTGGGGCTGGGAGGATGGCTCTGTCAACC 641
Db 241 TACCTTGGCCCTCTATGGCAATAGAGGCTTGGGGTGGGAGGATGGCTCTGTCAACC 300

Qy 642 CGCGCTCCCGCTAGTTGGGGCCCAACGACCCCGCGGTAGGTGCGTAACCTGGGT 701
Db 301 CGCGCTCCCGCTAGTTGGGGCCCAACGACCCCGCGGTAGGTGCGTAACCTGGGT 360

Qy 702 AAGGTATCATGATACCTTACATGCGGCTTCGCCGATCTCATGGGGTACATTCGCTCGTC 761
Db 361 AAGGTATCATGATACCTTACATGCGGCTTCGCCGATCTCATGGGGTACATTCGCTCGTC 420

Qy 762 GCGGCCCTAGGGGCGCTGCCAGGCGCTTGGCACACGGTGTCCGGGTTCTGGAGGAC 821
Db 421 GCGGCCCTAGGGGCGCTGCCAGGCGCTTGGCACACGGTGTCCGGGTTCTGGAGGAC 480

Qy 822 GCGGTGAACATATGCAACAGGGAACCTTGGCCGGTGTCTCTTTCTATCTTCTCTTTGGGT 881
Db 481 GCGGTGAACATATGCAACAGGGAACCTTGGCCGGTGTCTCTTTCTATCTTCTCTTTGGGT 540

Qy 882 CTGCTGTCTCTTGTGACCAATCCAGCTTCCGCTTATGAAGTGGCAACGTCGCGGATA 941
Db 541 CTGCTGTCTCTTGTGACCAATCCAGCTTCCGCTTATGAAGTGGCAACGTCGCGGATA 600

Qy 942 TACCATGTACGAAACGACGCTTCCAACTCAAGCATTTGTATGAGGACGCGGACGTCATC 1001
Db 601 TATCATGTACGAAACGACGCTTCCAACTCAAGCATTTGTATGAGGACGCGGACGTCATC 660

Qy 1002 ATGCATACCTCCGGGTGCGCTGCTGTTTACGAGGGTAAACAGTCTCCGTTCTCGGTA 1061
Db 661 ATGCATACCTCCGGGTGCGCTGCTGTTTACGAGGGTAAACAGTCTCCGTTCTCGGTA 720

Qy 1062 GCGCTCACTCCACGCTCGCGGCGAGGATGCCAGGTCCTCCACTAGCACAATACGACGC 1121
Db 721 GCGCTCACTCCACGCTCGCGGCGAGGATGCCAGGTCCTCCACTAGCACAATACGACGC 780

Qy 1122 CAGCTCGACTTCTGCTGGGAGCGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTC 1181
Db 781 CAGCTCGACTTCTGCTGGGAGCGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTC 840

Qy 1182 TGGGGATCTATTTTCTGCTCCAGCTGTTTCACTTCTCGCTCGCGGATGAGACA 1241
Db 841 TGGGGATCTATTTTCTGCTCCAGCTGTTTCACTTCTCGCTCGCGGATGAGACA 900

Qy 1242 GTGAGGACTGCAACTGCTCAATCTATCCGCGCATGTATCAGGTCAACCGATGGCTGG 1301
Db 901 TTACAGGACTGTAACCTGCTCAATTTATCCCGGCGCATGTGCGGTCACCGTATGGCTGG 960

Qy 1302 GATATGATGAACTGGGTCACTACAAAGCGCTAGTGTGCGAGTTGCTCCGATC 1361
Db 961 GACATGATGAACTGGGTCACTACAAAGCGCTAGTGTGCGAGTTGCTCCGATC 1020

Qy 1362 CCACAGCTGCTGAGACATGTTGGGGGCGCCACCTGGGGAGTCTCTGGGGGCTTTGCC 1421
Db 1021 CCACAGCTGCTGAGACATGTTGGGGGCGCCACCTGGGGAGTCTCTGGGGGCTTTGCC 1080

Qy 1422 TACTATTCATGTTAGGGAACCTGGGCTAAGGTTCTGATTTGGCGCTACTTTTCCGCGC 1481
Db 1081 TACTATTCATGTTAGGGAACCTGGGCTAAGGTTCTGATTTGGATGCTACTTTTCTGCGC 1140

Qy 1482 GTTGACGGGAGACCCACACGACGGGAGGGTGGCGGCGCACACCTCCGGGTTCAAG 1541
Db 1141 GTTGACGGGAGACCCACACGCTGACAGGGGGGCGCAAGCAAAACCAACAGGCTCGTG 1200

Qy 1542 TCCCTTTTCTCATCTGGGGCTCTCAGAAATCCAGCTTGTGAATACCAACGCGAGCTGG 1601
Db 1201 TCCATGTTGCAAGTGGGGCGCTCTCAGAAATCCAGCTTGTGAATACCAACGCGAGCTGG 1260

Qy 1602 CACATCAACAGGACTGCGCTAAATTTGCAATGACCTCCCTCCAAACTGGGTTCTTTGCCGCG 1661
Db 1261 CACATCAACAGGACTGCGCTAAATTTGCAATGACCTCTCTCCAGACTGGGTTCTTTGCCGCG 1320

Qy 1662 CTGTTTTACGCACAAAGTTCAACTGCTCGGGTGGCGGAGCGCATGGCCAGCTGCCGC 1721
Db 1321 CTGTTTTACGCACAAAGTTCAACTGCTCGGGTGGCGGAGCGCATGGCCAGCTGCCGC 1380

Qy 1722 CCCATTGATGGTTGGCCAGGGTGGGGCCCCCATCACTATATAGCCCTAAACAGCTCG 1781
Db 1381 ACCATTGCAAGTTGCAACAGGATGGGGTCCCCATTACTTATGTGAGTCTAGCAGATCA 1440

Qy 1782 GATCAGAGGCTTATTTGCTGGCATTTAGCGGCTCGACCGTGTGGTGTGCTACCCCGCTCG 1841
Db 1441 GACACAGAGGCTTATTTGCTGGCATTTAGCGGCTCGACCCACTCCCAATGTACCATGCTCGCTCG 1500

Qy 1842 CAGGTGTGTGCTCCAGTGTATTGTTTTCACCCCAAGCCCTGTGTGGTGGGGACCAACCGAT 1901

QY 4062 CCGGCTGCGTATGACGCCAAGGTTACAAAGGTGCTCGTCTGAACCCGTCGGTTCGCCGC 4121
Db 3721 CCGGCTGCATATGCAGCCCAAGGTTACAAAGGTGCTCGTCTCAATCCGTCGGTTCGCCGC 3780
QY 4122 ACCTTAGGTTTGGGCGTATATGTCCAAAGGCAACGGGTATGACCCCTTAACATCAGAACT 4181
Db 3781 ACCTTAGGTTTGGGCGTATATGTCTAAGGCACACGGTATTGACCCCAACATCAGAACT 3840
QY 4182 GGGTTAAGGACCAATTACACGGGCGGCTCCATTAAGTACTCCACCTATGGCAAGTTCCTT 4241
Db 3841 GGGTTAAGGACCAATTACACGGGCGGCTCCATTAAGTACTCCACCTATGGCAAGTTCCTT 3900
QY 4242 GCCGACGGTGGCTGTTCTGGGCGGCTATGACATCATATAATATGATAGTGCCACTCA 4301
Db 3901 GCCGATGGTGGTCTCTGGGCGGCTTATGACATCATATAATATGATAGTGCCACTCA 3960
QY 4302 ACTGACTCGACTACCACTTTGGGCGCATCGGCACAGTCTCTGGACCAAGCGGAGCGGTGGA 4361
Db 3961 ACTGACTCGACTACCACTTTGGGCGCATCGGCACAGTCTCTGGACCAAGCGGAGCGGTGGA 4020
QY 4362 GCGGCGTCTGTCGTGTCGCCACCGGTACACCTCCGGGATCGGTATACCGTGCCACACCC 4421
Db 4021 GCGGCGTCTGTCGTGTCGCCACCGGTACACCTCCGGGATCGGTATACCGTGCCACACCC 4080
QY 4422 AATATCGAGAAATAGGCTCTGTCACCAATAGGAGATCCCTTCTATGGCAAGCCATC 4481
Db 4081 AACATCGAGAGGTGGCCCTGTCTAATACTGGAGAGATCCCTTCTATGGCAAGCCATC 4140
QY 4482 CCCATTGAGGCCATCAAGGGGGGAGGCATCTCAATTTCTGCCATCCCAAGAGAAATGT 4541
Db 4141 CCCATTGAGGCCATCAAGGGGGGAGGCATCTCAATTTCTGCCATCCCAAGAGAAATGT 4200
QY 4542 GACGAGTCCGCGCAAGCTGACAGGCTCGGACTGAACGCTGTAGCATATTAACCGGGC 4601
Db 4201 GACGAGTCCGCGCAAGCTGTCAGGCTCGGATCAACGCTGTGGGTATTACCGGGG 4260
QY 4602 CTTGATGTCCGTCATACCGCTTATCGGAGAGTCTGTTGTGTTGGCAACAGAGCTCTA 4661
Db 4261 CTCGATGTCCGTCATACCAACTATCGGAGAGTCTGTTGTGTTGGCAACAGAGCTCTG 4320
QY 4662 ATGACGGGTTTACCGGGGATTTGACTCAGTGATCACTGCAATATACATGTCACCCAG 4721
Db 4321 ATGACGGGCTATACGGGCGACTTTGACTCAGTGATCACTGTAACACATGTGTCAACCCAG 4380
QY 4722 ACAGTCAGTTCAGTTGGATCCCACTTACCAATTGAGACGACCGTSCCCCAAGAC 4781
Db 4381 ACAGTCAGTTCAGTTGGATCCCACTTACCAATTGAGACGACCGTSCCTCAAGAC 4440
QY 4782 GCGGTGTCGCGCTCGCAACGCGGAGGTAGAACTGGCAGGGGTAGGATGGCATCTACAGG 4841
Db 4441 GCAGTGTGCGCTCTGCAGCGCGGGGTAGGACTGGCAGGGGTAGGAGGCACTACAGG 4500
QY 4842 TTTGTGACTCCAGGAGAACCGCCCTCGGGCATGTTTCGATTCCTCGGTCTGTGTGAGTGC 4901
Db 4501 TTTGTGACTCCGGGAGAACCGCCCTCGGGCATGTTTCGATTCCTCGGTCTGTGTGAGTGC 4560
QY 4902 TATACGCGGCGTGTGCTTGTGTATAGCTACGCGCGCTGAGACCTCGGTTAGGTTGCGG 4961
Db 4561 TATACGCGGCGTGTGCTTGTGTATAGCTACGCGCGCTGAGACCTCGGTTAGGTTGCGG 4620
QY 4962 GCTTACCTTAATACACAGGGTTGCCCTCTGCCAGGACCATCTGGAGTCTTGGGAGAGC 5021
Db 4621 GCCTACCTTGAACACACAGGGTTGCCCTTGGCAGGACCACTTGGAGTCTTGGGAGAGT 4680
QY 5022 GTCTTACAGGCTTCAACCAATAGATGCCAATTCCTGTCCAGACTTAAACAGGACGGA 5081
Db 4681 GTCTTACAGGCTTCAACCAATAGATGCCAATTCCTGTCCAGACTTAAACAGGACGGA 4740
QY 5082 GACAACTTTCTTACTCTGGTGCATATCAAGCTACAGTGTGGCGGAGGGCTCAAGCTCCA 5141
Db 4741 GACAACTTTCTTACTCTGGTGCATATCAAGCTACAGTGTGGCGGAGGGCTCAGGCCCA 4800

QY 5142 CTTCCATCTGTGGGACCAATGTGGAAAGTGTCTCATACGGCTGAAACCTACACTGCACGGG 5201
Db 4801 CTTCCATCTGTGGGACCAATGTGGAAAGTGTCTCATACGGCTGAAACCTACACTGCACGGG 4860
QY 5202 CCAACACCCCTCTGTATAGGCTAGAGCCGTCCAAATAGAGTCACTCTCACACACCC 5261
Db 4861 CCAACACCCCTTCTGTATAGGCTAGAGCCGTCCAAATAGAGTCACTCTCACACACCC 4920
QY 5262 ATAACATAATACATCATGSCATGTCGCTGAGCTCGGAGGTGTCTATAGCACTGG 5321
Db 4921 ATAACATAATACATCATGSCATGTCGCTGAGTCTCGAGTCTGCTACCTAGCACTGG 4980
QY 5322 GTGCTGTAGGCGGAGTCTCTGAGCTTTGGCGGCATATGCTGTAGCAGCAGCAGTGTG 5381
Db 4981 GTGCTGTGGGCGGAGTCTCTGAGCTCTGGCGCGTATTGCTGTGCAACAGGCACTGTG 5040
QY 5382 GTCATTTGTGGGAGGATCATCTTTGTCGGGAGCAGCTGTCTTCCGCACAGGGAAGTC 5441
Db 5041 GTCATTTGTGGGTAGGATTAATCTTTGTCGGGAGCGCGCCATTTGTTCCGCACAGGAGCTT 5100
QY 5442 CTCCTACAGAGTTCGATGAGATGGAAGAGTGTGCTCTCAAACTTCTTACATCTGAGCAG 5501
Db 5101 CTCCTACAGAGTTCGATGGAATGGAAGAGTGTGCTCTCGCACTCTCTTACATCTGAGCAG 5160
QY 5502 GGAATGAGCTCGCCGAGCAATTCAAAGCAAAAGCGCTCGGGTTGTTGCAAAAGCGCACCC 5561
Db 5161 GGAATGAGCTCGCCGAGCAATTCAAAGCAAAAGCGCTCGGGTTAATGTCAAACAGCCACC 5220
QY 5562 AAGCAAGCGGAGGCTGCTCTCCGTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACCTTC 5621
Db 5221 AACAAGCGGAGGCTGCTCTCCGTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACANTC 5280
QY 5622 TGGCGCAAGCACATGTGGAAATTCATCAGCGGAATACAGTACCTAGCAGCTTATCCACT 5681
Db 5281 TGGCGCAAGCACATGTGGAAATTCATCAGCGGAATACAGTACCTAGCAGCTTATCCACT 5340
QY 5682 CTCCTTGGAAACCCCGATAGCATCATTTGATGGCATTTACAGCTTCTATCCTAGCCTCG 5741
Db 5341 CTCCTTGGAAACCCCGATAGCATCATTTGATGGCATTTACAGCTTCTATCAGCCTCTATC 5400
QY 5742 CTCACCAACCAAAACCCCTCTCTTTAAATCTTGGGGGATGGGTGGCTGCCAACTC 5801
Db 5401 CTCACCAACCAAAAGTACCTCTCTTTAAATCATCTTGGGGGGTGGGTGGCTGCCAACTC 5460
QY 5802 GCTCTCTCCAGCGCTGCTGAGCTTTCTGGGCGCGGCATCGCCGAGCGGCTGTGGC 5861
Db 5461 GCGCCCGCAGCGCTGCTGAGCTTTCTGGGCGCGGCATCGCCGCTGCGGCTGTGGC 5520
QY 5862 AGCATAGGCTTGGGAGGTTGCTCGTGAGCATCTTGGCGGGCTATGGCGAGGGGTAGCC 5921
Db 5521 AGCATAGGCTTGGGAGGTTGCTGTGGACATTTCTGGCGGGTTATGGAGCAGGAGTGGCC 5580
QY 5922 GCGGCACTCTGCGCTTTTAAAGGTTCATCAGCGCGGAGGTGCCCTCCACGAGGACCTGGTC 5981
Db 5581 GCGGCGCTCTGCGGCTTTTAAAGGTTCATGAGCGCGGAGATGCCCTCCAAGGAGACTGGTC 5640
QY 5982 AACTTACTCTCGCATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6041
Db 5641 AACTTACTCTCTGCT 5700
QY 6042 ATACTGCTCGGACGTTGGGCGCGGAGGCGCTGTGCAAGTGAACCCGCTGATA 6101
Db 5701 ATACTGCTCGGACGTTGGGCTCGGGAGGCGGCTGTGCAAGTGAACCCGCTGATA 5760
QY 6102 GCGTTCGCTCGCGGGTAAACAGCTCTCCCTACGACATATGTGCTGAGGAGCGCT 6161
Db 5761 GCGTTCGCTCGCGGGTAAATCATGTTTCCCGCCGCACTATGTGCTGAGGAGCGCC 5820
QY 6162 GCACGAGTGTCACTCAGATCTCTTAGCTTACCACTCACTCACTCACTCACTCACTCACTCACT 6221
Db 5821 GCACGAGTGTCACTCAGATCTCTCAGATCTCTCAGATCTCTCAGATCTCTCAGATCTCT 5880
QY 6222 CACCAAGTGAATTAATGAGGACTGCTCTACGCCATGCTCTCGGCTCTGCTGCTGCTGCTGCT 6281

||||| 5881 CACCAGTGGATTAATGAAGACTGCTCCACACCGTGTTCGGCTCGTGGCTAAGGATGTT 5940
||||| 6282 TGGGATTTGGATATGACCGGTGTGACTGACTTCAAGACCTGGCTCCAGTCCAAATCTCTGT 6341
||||| 5941 TGGGACTGGATATGACCGGTGTGACTGACTTCAAGACCTGGCTCCAGTCCAAATCTCTGT 6000
||||| 6342 CCGCGGTTACCGGGAGTCCCTTTCTCTGTATGCCAAACCGGGTACAAAGGAGTCTGGCGG 6401
||||| 6001 CCGCAGCTACCTGGAGTCCCTTTTCTCTGTGCCAAACCGGGTACAAAGGAGTCTGGCGG 6060
||||| 6402 GGGGAGCGCATCATGCAAAACCACTGCTCCCATCGGAGACAGATCGCCGGAATGTCAAA 6461
||||| 6061 GGAGACGGCATCATGCAAAACCACTGCCCATGTGGAGACAGATCACCGGAATGTCAAA 6120
||||| 6462 AACGGTTCCATCAGAGTCTAGGGCTAGAACCTGCAGCAACACGTGGCGACGGAAGTTC 6521
||||| 6121 AACGGTTCCATCAGAGTCTAGGGCTAGAACCTGCAGCAACACGTGGCGATGGAACATTC 6180
||||| 6522 CCATCAACCGCATACACACGGGACCTTTGCACACCTCTCCCGCGGCCCAACTATTTCCAGG 6581
||||| 6181 CCATCAACGCATACACACGGGCCCTTGCAACCTCTCCAGCGGCAAACTATTTCTAGG 6240
||||| 6582 GGGCTATGCGGGTGGCTGCTGAGGAGTACGTGGAGTTACCGGTGTGGGGATTTCCAC 6641
||||| 6241 GCGCTGTGGCGGTGGCGCTGAGGAGTACGTGGAGTCAACCGGGTGGGGATTTCCAC 6300
||||| 6642 TACGTGACGGGCATGACCACTGACAACTGACAACTGACAACTGACAACTGACAACTGACAA 6701
||||| 6301 TACGTGACGGGCATGACCACTGACAACTGACAACTGACAACTGACAACTGACAACTGACAA 6360
||||| 6702 TTCTTTCACGGAGTGGATGGAGTGGCGTTGACAGGTACGCTCCGCGGTGCAAACTCTTT 6761
||||| 6361 TTCTTCTCGGAGTGGACGGAGTGGCGTTGACAGGTACGCTCCGCGGTGACAGGCTCTC 6420
||||| 6762 CTACGGGAGGAGTGGATTTCCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 6821
||||| 6421 CTACGGGAGGAGTGGATTTCCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTC 6480
||||| 6822 CCATGGAGCGCAACCGGACGTACAGTGTCTTACGTCTTACGTCTTACGTCTTACGTCTTACGTCT 6881
||||| 6481 CCATGGAGCGCAACCGGATGACGTGTCTTACGTCTTACGTCTTACGTCTTACGTCTTACGTCT 6540
||||| 6882 ATTACAGCAGACGCGCTAAGCGTAGGCTGGCTAGAGGCTCTCCCGCTCTTTAGCCAGC 6941
||||| 6541 ATCAGCAGAAACGGCTAAGCGTAGGTTGGCCAGGGGGTCTCCCGCTCTTTGGCCAGC 6600
||||| 6942 TCATCAGCTAGCCAGTGTCTCGGCTTCTTTGAAGGGACATGCACTACCCACCATGAC 7001
||||| 6601 TCTTCAGCTAGCCAGTGTCTCGGCTTCTTTGAAGGGACATGCACTACCCACCATGTC 6660
||||| 7002 TCCCGGAGCGTGCATCATCAGGCGCAACCTCTTGTGGCGGAGAGATGGCGGGAAC 7061
||||| 6661 TCTCCGAGCGCTGACCTCATCAGGCGCAACCTCTTGTGGCGGAGAGATGGCGGGAAC 6720
||||| 7062 ATCAGCTCGGTGGAGTACAGAAATAGGTAGTAAATTTCTGAGTCTTTGAAACCGCTTCAAC 7121
||||| 6721 ATCAGCGGTGGAGTTCGGAGAAAGGTGGTGGTCTGAGTCTTTCAGCCCGCTTCA 6780
||||| 7122 GCGGAGGGGATCAGAGGAGATATCCGTTCGGCGGAGATCTCGGAAATCCAGGAAG 7181
||||| 6781 GCGGAGGAGATCAGAGGAGATATCCGTTCGGCGGAGATCTCGGAAATCCAGGAAG 6840
||||| 7182 TTCCCGTACGCTTGGCCATATGGGACCGCCGGAATCAATCTTCCACTGTAGATGCC 7241
||||| 6841 TTCCCGGAGCATGCGCATCTGGCGCGCCCGGATTAACCCCTCCACTGTAGAGTCC 6900
||||| 7242 TGAAGGACCCGAGCTACGCTCTCCGCTGTACACGATGCGCATTTGCCACTTACCAAG 7301
||||| 6901 TGAAGGACCCGAGCTACGCTCTCCGCTGTGACCGGGTGGCGGTGGCCACTTATCAAG 6960
||||| 7302 GCTCTCTCAATACCACTCCAGGAAAGAGGACGGTTGTCTGACAGAAATCCAAATGTG 7361
|||||

Db 6961 GCCCTTCAATACCACTCCACGGAGAAAGGACGGTTGTCTTAACAGAGTCTCCCGTG 7020
QY 7362 TCTTCTGCTTTGGGGAGCTCGCCACTAAGACCTTCGCTAGCTCGGATCTCGGCGCGTT 7421
Db 7021 TCTTCTGCTTTAGCGGAGCTCGCTACTAAGACCTTCGCGAGCTCCGAATCATCGCGCGTC 7080
QY 7422 GATAGCGGCA CGGCGACCGCCCTTCTGACTCTGCGCTCCGACGACGGTGACAAAGATCC 7481
Db 7081 GACAGCGCACGGCGACCGCCCTTCTGACACGAGCTCCGACGAGTGAACAAAGATCC 7140
QY 7482 GACGTGTAGTGTACTCTCTCATGCCCCCTTGAAGGGAGCGGGGACCCCATCTC 7541
Db 7141 GACGTGTAGTGTACTCTCTCATGCCCCCTTGAAGGGAGAACCCGGGAGACCCCATCTC 7200
QY 7542 AGCAGCGGTCTTGGTGTACCGTGTAGTACGAGGCTAGTGTAGGATGTGTCTGTCTGTCTCA 7601
Db 7201 AGTGTGCGGTCTTGGTGTACCGTGTAGTGTAGGATGTGTCTGTCTGTCTCA 7260
QY 7602 ATGTCTTATACGTGTGACAGGCGCTGTATCATCGCATCGCTGTGGAGGAAAGTAAAGTGT 7661
Db 7261 ATGTCTTACATGTGACAGGCGCTTGTATCATCGCATCGCTGTGGAGGAAAGTAAAGTGT 7320
QY 7662 CCATCAACCGGTGTAGCAACTCTTTGTCTGTGTACCAACATGTCTACGCCACAACA 7721
Db 7321 CCCATCAACCGGTGTAGCAACTCTTTGTCTGTGTACCAACATGTCTTATGCCACAACA 7380
QY 7722 TCCGACGCGAAGCTCCGCGCAGAAAGGTTCACCTTTGACAGATTTGCAAGTCTCTGGAT 7781
Db 7381 TCTGTGACGCGAGCTCGCGCAGAAAGGTTCACCTTTGACAGATTTGCAAGTCTCTGGAC 7440
QY 7782 GATCATTAACGGGACGTACTCAAGGAGATGAAGGCGAAGGCGTCCACAGTTAAGGCTAAG 7841
Db 7441 GACCATAACGGGACGTCTCAAGGAGATGAAGGCGAAGGCGTCCACAGTTAAGGCTAAG 7500
QY 7842 CTTCTATCTATAGAGGCGCTGTCAAGTGTAGCGCCCCACATTTGCGGCCAAATCCAAATTT 7901
Db 7501 CTCTTATCCGTAGAGGAAAGCTGCAAGCTGACGCCCCCCACATTTGCGGCCAAATCCAAATTT 7560
QY 7902 GGTATGGGCAAAAGGAGCTCCGGAACCTATCCAGCAGGCGGCTTAAACACATCCGCTCC 7961
Db 7561 GGTATGGGCAAAAGGAGCTCCGGAACCTATCCAGCAGGCGGCTTAAACACATCCGCTCC 7620
QY 7962 GTGTGGGAGGACTTGTCTGGAAGACACTGAAACACCAATTTGACACCACTATCGCAAAA 8021
Db 7621 GTGTGGGAGGACTTGTCTGGAAGACACTGTCGACCAATTTGACACCACTATCGCAAAA 7680
QY 8022 AGTGTGTTTCTGCGTCCAAACAGAGAGGAGGCGCGAAGCAGCTCGCCTTATCGTA 8081
Db 7681 AATGAGGTTTCTGTTGTCCAAACAGAGAAAGGAGCGGTAAGCAGCGCCCTTATCGTA 7740
QY 8082 TTCCAGACCTGGGAGTTCTGTATGCGAGAGATGGGCCCTTTAGAGTGTCTCCACC 8141
Db 7741 TTCCAGATCTGGGAGTCCGTGTATGCGAGAAAGTGGGCCCTTATGATGTGTCTCCACC 7800
QY 8142 CTTCTCAGGCGGTGATGGGCTCTCTCATACGATTTCAATACTCCCCCAAGCAGCGGGTTC 8201
Db 7801 CTTCTCAGGCTGTGATGGGCTCTCTCATACGATTTCAAGTCTCTCTGGGAGCGAGTTC 7860
QY 8202 GAGTTCTGGTGAATACCTGGAAATCAAGAAATGCCCTATGGGCTTCTCATATGACACC 8261
Db 7861 GAGTTCTGGTGAATACCTGGAAATCAAGAAATCAAGAAATCAAGAAATCAAGAAATCAAGAAAT 7920
QY 8262 CGCTGTTTTGTACTCAACGGTCACTGAGATGACATTCGTTGTGAGAGTCAATTTACCAA 8321
Db 7921 CGCTGTTTTGTACTCAACGGTCAACCGGACCGGACATTCGTTGTGAGAGTCAATTTACCAA 7980
QY 8322 TGTGTGACTTGGCCCCCGGAGCAGACAGGCGCAATAGGTCGCTCACAGAGCGGCTTTAC 8381
Db 7981 TGTGTGACTTGGCCCCCGGAGCAGACAGGCGCAATAGGTCGCTCACAGAGCGGCTTTAT 8040
QY 8382 ATCGGGGTTCCTGTACTTAACTCAAAAGGAGAACTGCGGTTATCGCGGTGCGCGCA 8441
Db 8041 ATCGGGGTTCCTGTACTTAACTTCAAAAGGAGAACTGCGGTTATCGCGGTGCGCGCG 8100

QY	8442	AGTGGCGTCTGACGACTAGCTGCGTAATACCTCACATGTTACTTGAAGGCCACTGCA	8501
Db	8101	AGCGGCGTCTGACGACTAGCTGCGTAACACCCCTCACATGTTACTTGAAGGCCCTCTGCA	8160
QY	8502	GCCTGTGAGCTGCAAGCTCCAGGACTGCAGATGCTGTAACGGAGACGACCTTGTG	8561
Db	8161	GCCTGTGAGCTGCAAGCTCCAGGACTGCAGATGCTGTAACGGAGACGACCTCGTC	8220
QY	8562	GTTATCTGTGAAGCCGCGGACCCAGGAGATGCGCGGCCCTACGAGCCTTCACGGAG	8621
Db	8221	GTTATCTGTGAAGCCGCGGACCCAGGAGATGCGCGGCCCTACGAGCCTTCACGGAG	8280
QY	8622	GCTATGACTAGGTATTCCGCCCCCCCCGGGGATCCGCCCCAACAGAAATACGACCTGGAG	8681
Db	8281	GCTATGACTAGGTACTTCGCCCCCCCCGGGGACCCGCCCCAACAGAAATACGACCTGGAG	8340
QY	8682	CTGATAACATCATGTTCTCTCAATGTGTAGTGGCGACGATGCATCTGSCAAAAGGTA	8741
Db	8341	CTGATAACATCATGTTCTCTCAATGTGTAGTGGCGACGATGCATCAGGCAAAAGGGTG	8400
QY	8742	TACTACTCACCGCTGACCCACCCCTTTCAGCGGCTGCTGGGAGACAGCTAGA	8801
Db	8401	TACTACTCACCGCTGATCCACCCCTTAGCAGGCTGCTGGGAGACAGCTAGA	8460
QY	8802	CACACTCCAATCAACTCTTGCTAGGCAATATCATCATGTATGCGCCACCTATGGGCA	8861
Db	8461	CACACTCCAAGTTAACTCTTGCTAGGCAATATCATCATGTATGCGCCACCTATGGGCA	8520
QY	8862	AGGATGATCTGATGACTCTCTTTCTCCATCTCTAGCTCAAGAGCAACTTGAATAA	8921
Db	8521	AGGATGATCTGATGACTCTCTCTCTCCATCTCTAGCTCAAGAGCAACTTGAATAA	8580
QY	8922	GCCTGGAATGTCAGATCTAGGGGCTGTACTCCATTTAGGCACCTTGACCTACCTCAG	8981
Db	8581	GCCTGGAATGTCAGATCTAGGGGCTGTACTCCATTTAGGCACCTTGACCTACCTCAG	8640
QY	8982	ATCATTCAGACCTCCATGCTTTAGCGCATTTACACTCCACAGTTACTCTCCAGGTGAG	9041
Db	8641	ATCATTCAGACCTCCATGCGCTTTAGCGCATTTACACTCCATAGTTACTCTCCAGGTGAG	8700
QY	9042	ATCAATAGGTGGCTTCATGCTCAGGAACTTGGGGTACACCCCTTGGGAACTGGAGA	9101
Db	8701	ATCAATAGGTGGCTTCATGCTCAGGAACTTGGGGTACACCCCTTGGGAGTCTGGAGA	8760
QY	9102	CATCGGCGCAGAGTGTCCGGGCTAAGCTACTGTCCAGGGGGAGGGCCGCCACTTGT	9161
Db	8761	CATCGGCGCAGAGGCTCCGGCTAGGCTACTGTCCAGGGAGGGGGCGGCCACTTGT	8820
QY	9162	GGCAGATACCTCTTAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCCAATCCCGCC	9221
Db	8821	GGCAATACCTCTTCAACTGGGCACTAAGGACCAAGCTTAAACTCACTCCCAATCCCGCT	8880
QY	9222	CGTCCCAAGCTGGACTTGTCTGGCTGGTTCCGCTGGTTACAGCGGGGAGACATATAT	9281
Db	8881	CGTCCCGGCTGGACTTGTCCGGCTGGTTCCGCTGGTTACAGCGGGGAGACATATAT	8940
QY	9282	CACAGCCTGTCTCGTGGCCGACCCCGCTGGTTCCGTTGTCCTACTCTTCTGTGA	9341
Db	8941	CACAGCCTGTCTCGTGGCCGACCCCGTTGGTTTCATGCTGTGCTACTCTACTTCTGTA	9000
QY	9342	GGGTAGGCATTTACTGCTCCCAACCGA	9371
Db	9001	GGGTAGGCATTTACTGCTCCCAACCGA	9030

This Page Blank (uspto)

Db 242 CCGAGTAGTGTTGGGTCGCCGAA

[illegible]

3542	Db	GA	CTGCGTCAATGGCGGTGTGGACCGTCTACCAATGGTGGCGGCTCGAAGACCCTGGC	3601
3614	Qy	CGGTCCAAAAGGTCCAAATCACCCAAAATGTACACCAATGTAGACCTTGGACCTCGTGGGCTG	3673	
3602	Db	CGGCCGAAGGTCCTCAATCACCCAAATGTACACCAATGTAGACCAAGGACTCGTGGGCTG	3661	
3674	Qy	GCAGGCGCCCCCGGGCGGGCTCCATGACACCATGCACTGCACTGTGGCAGCTCCGACCTTTA	3733	
3662	Db	CGCGGCGCCCCCGGGCGGGCTCCATGACACCGTGCACCTCGGGGAGCTCGGACCTTTA	3721	
3734	Qy	CTTGGTCAAGAGACATGCTCATGTTCATTCGGGTGCGCGCGAGAGGCGACAGCAGGGGAAG	3793	
3722	Db	CTTGGTCAAGAGATGCTCATGTTCATTCGGGTGCGCGCGAGAGGCGACAGCAGGGGAAG	3781	
3794	Qy	TCTACTCTCCCCAGGCCCTCTCTACCTGAAAGGCTCTCGGGGTGCTCCATGCTTTG	3853	
3782	Db	CCTGCTTTCCCCAGGCCCATCTCTACCTGAAAGGCTCTCTCGGGTGACCACTGCTTTG	3841	
3854	Qy	CCCTTCGGGGACAGTGTGGGGCTCTTCGGGGTGTGTGTGCACCCGGGGGTTCGGAA	3913	
3842	Db	CCCTTCGGGGACAGTGTGAGGCATCTTTCGGGGTGTGTGTGCACCCGGGGGTTCGGAA	3901	
3914	Qy	GGCGGTGCACTTTCATCCCGTTGAGTCTATGGAACCTACCATCGGGTCTCCGGTCTTCAC	3973	
3902	Db	GGCGGTGCACTTTCATCCCGTTGAGTCTATGGAACCTACCATCGGGTCTCCGGTCTTCAC	3961	
3974	Qy	AGACAACTCAACCCCCCGGCTGTACCGCAGACATTCCAAAGTGGCACATCTGCACGCTCC	4033	
3962	Db	AGACAACTCATCCCTTCGGCGGTACCGCAAAATTCCAAAGTGGCACATTTACACGCTCC	4021	
4034	Qy	TACTGGCAGCGGCAAGAGACCAAGTGC CGGCTGCGTATGCAGCCCAAGGTACAGGT	4093	
4022	Db	CACCTGGCAGCGGCAAGAGACCAAGTGC CGGCTGTCATATGCAGCCCAAGGTACAGGT	4081	
4094	Qy	GCTCGTCTGAAACCGGTCCGTTGCGGCCACTTAGGGTTTGGGCGGTATATGTCAAAGGC	4153	
4082	Db	GCTCGTCTTAAACCGGTCCGTTGCGGCCACATTTGGGCTTTGGAGCGGTATATGTCAAAGGC	4141	
4154	Qy	ACCGGTATCGACCCCTAACATCAGAACTGGGGTAAGGACCATTAACACGGCGGCTCCAT	4213	
4142	Db	ACATGGCATCGAGCCTTAACATCAGAACTGGGGTAAGGACCATTAACACGGCGGCTCCAT	4201	
4214	Qy	TACGTATCCACCTATGGCAAGTTCCTTGCAGCGGTGGCTGTCTCGGGGGCGCTATGA	4273	
4202	Db	CACGTACTCCACCTATTTGCAAGTTCTTTCGCGACGGTGGATGTCTCCGGGGCGCTATGA	4261	
4274	Qy	CATCATATATGTGATGAGTGCCACTCAACTGACTCGACTTACCATTCTTGGGCATCGGCAC	4333	
4262	Db	CATCATATATGTGATGATGCCACTCAACTGACTCGACTTACCATTCTTGGGCATCGGCAC	4321	
4334	Qy	AGTCTCGACCAAGCGGAGACGGCTGGAGCGCGCTCGTGTGTTCGCGCACCGCTACACC	4393	
4322	Db	AGTCTCGATCAGGCAGAGACGGCTGGAGCGCGCTCGTGTGTTCGCGCACCGCTACACC	4381	
4394	Qy	TCCGGGATCGGTACCGTGCCACACCCCAATATCGAGGAATAGCCCTGTCCACAAATGG	4453	
4382	Db	TCCGGGATCGATCACCGTGCCACACCCCAATATCGAGGAAGTGGCCCTGTCCAACTGG	4441	
4454	Qy	AGAGATCCCTCTATGTGCAAGCCATCCCATTTAGGCCATCAAGGGGGGAGGCATCT	4513	
4442	Db	AGAGATCCCTCTATGTGCAAGCCATCCCATTTAGGCCATCAAGGGGGGAGGCATCT	4501	
4514	Qy	CATTTCTGCCATTCGAAGAAATGTGACGAGCTCGCGCAAGAGTGCAGACGCTCGG	4573	
4502	Db	CATCTTCTGCCATTCGAAGAAATGTGACGAGCTCGCGCAAGAGTGCAGACGCTCGG	4561	
4574	Qy	ACTGAACGCTGAGCATATTAACCGGGGCTTGTATGTGTCCGTATACCGCTATTCGAGA	4633	
4562	Db	ACTCAATGCTGTAGCGTATTACCGGGGTCTCGATGTGTCCGTCTATACCGACTAGCGGAGA	4621	
4634	Qy	CGTGTGTGCGGGCAAGACGCTCTAATGACGGGTTTTCACCGCGATTTTGTACTCAGT	4693	
4622	Db	CGTGTGTGCGGGCAAGACGCTCTAATGACGGGTTTTCACCGCGACTTTTGTACTCAGT	4681	

QY 4694 GATCGACTGCAATACATGTGTCAACCAGACAGTGCAGTTTCAGCTTGGATCCACACTTCAC 4753
DB 4682 GATCGACTGCAACACATGTGTCAACCAGACAGTGCAGTTTCAGCTTGGATCCACACTTCAC 4741
QY 4754 CATTGAGACGACGACCGTCCCAAGACGCGGTGTCGCGCTCGCAACCGGAGGTAGAAC 4813
DB 4742 CATTGAGACGACGACCGTCCCAAGACGCGGTGTCGCGCTCGCAACCGGAGGTAGAAC 4801
QY 4814 TGGCAGGGGTAGGAGTGGCATCTACAGTTTGTGACTCCAGGAGAACCGCCCTCGGGCAT 4873
DB 4802 TGGCAGGGGTAGGAGTGGCATCTACAGTTTGTGACTCCAGGAGAACCGCCCTCAGGCAT 4861
QY 4874 GTTTCGATTTCTTCGGTCTCTGTGTGAGTGTATGACGCGGCTGTGTCTGTGTATGAGCTCAC 4933
DB 4862 GTTTCGACTCTCTCGTCTCTGTGTGAGTGTATGACGCGAGGTGCGTCTGTGTATGAGCTCAC 4921
QY 4934 GCCCGCTGAGACCTCGGTTAGGTTTCGGGCTTACCTAAATACACAGGTTTGCCTGTG 4993
DB 4922 GCCCGCTGAGACCTCGGTTAGGTTTCGGGCTTACCTAAATACACAGGTTTGCCTGTG 4981
QY 4994 CCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGCTCACCACATAGATGCCCA 5053
DB 4982 CCAGGACCATCTAGAGTTCTGGGAGAGCGTCTTACAGGCTCACCACATAGATGCCCA 5041
QY 5054 CTTCTGTCTCCAGACTAAACAGCGAGGAGACAACTTCTTACCTGTGGCATATCAAGC 5113
DB 5042 CTTCTGTCTCCAGACCAACAGCGAGGAGAACCTCCCTACCTGGTAGCATCAAGC 5101
QY 5114 TACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAAATGTGGAAGTGTCT 5173
DB 5102 CACAGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGACCAAAATGTGGAAGTGTCT 5161
QY 5174 CATACGGCTGAACCTACATCTGACGCGGCAACACGCCCTGTGTATAGCTTAGGAGCGGT 5233
DB 5162 CATACGGCTGAACGCGCAACATCTGATGGGCAACGCGCTGTGTATAGCTTAGGAGCGGT 5221
QY 5234 CCAAAATGAGGTCACTCTCACACACCCATACTAAATACATCATGGCATGATCGCGC 5293
DB 5222 TCAAAATGAGGTCACTCTCACACACCCATACTAAATACATCATGGCATGATCGCGC 5281
QY 5294 TGACCTGGAGGTGCTCACTAGCACCTGGGTGCTGGTAGGCGGAGTCTTTGACAGCTTTGGC 5353
DB 5282 TGACCTGGAGGTGCTCACTAGCACCTGGGTGCTGGTAGGCGGAGTCTTTGCGGCTCTGGC 5341
QY 5354 CGCATACTGCTGACGACGAGGAGTGTGTCTATTTGTGGGACGATCATCTTGTCCGGGAA 5413
DB 5342 CGCGTACTGCTGACGACGAGGAGGTGTGTCTATTTGTGGGACGATCATCTTGTCCGGGAG 5401
QY 5414 GCAGTGTGCTTCCCGACAGGGAAGTCTCTTACAGGAGTTTCGATGAGATGGAAGAGTG 5473
DB 5402 GCAGTGTGTTATCCCGACAGGGAAGTCTCTTACAGGAGTTTCGATGAGATGGAAGAGTG 5461
QY 5474 TGCCTCACAACTTCTTATCATCGACGAGGAAATGCAGCTCGCGAGCAATTCAGACAAA 5533
DB 5462 TGCCTCACAACTTCTTATCATCGACGAGGAAATGCAGCTCGCGAGCAATTCAGACAAA 5521
QY 5534 GCGCTCGGGTTGTGCAAAACGGCCACGAAAGCGGAGGTGTGTCTCCCGTGGTGA 5593
DB 5522 GCGCTCGGATTTGTGCAAAACAGCCACCAAGCAAGCGGAGGTGTGTCTCCCGTGGTGA 5581
QY 5594 GTTCCAAGTGGCAGCCCTTGAGACTTCTTGGCGGAGCACATGTGGAATTTTCATCAGCGG 5653
DB 5582 GTTCCAAGTGGCAGCCCTTGAGGTTCTTGGCGGAGAACACATGTGGAATTTTCATCAGCGG 5641
QY 5654 AATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAAACCCCGGATAGCATCATTTGAT 5713
DB 5642 GATACAGTACTTGGCAGGCTTATCCACTCTGCTGGAAACCCCGGATAGCATCATTTGAT 5701
QY 5714 GGCATTTACAGCTTCTATCATAGCCGCTACACCCGAAACACCCCTCTCTTTAAACAT 5773
DB 5702 GGCATTTACAGCTTCTATCACCAGCCGCTCACCACCCAAAATACCCCTCTCTTTAAACAT 5761

QY 5774 CTTGGGGGATGGGTGGTGCCTCCCAACTCGCTCTCTCCAGCGCTGCGTCAGCTTTCTGGG 5833
DB 5762 CTTGGGGGATGGGTGGTGCCTCCCAACTCGCTCTCCCGCAGCGCTGCTTCGGCTTTCTGGG 5821
QY 5834 GCGCGGATCCCGGAGGCGGTGTGGAGCATATAGGCTTGGGAAGGTGCTCTGTGGACAT 5893
DB 5822 GCGCGGATTTGCGCGGTGCGCGGTGTGGCAGCATATAGGCTCTCGGAAGGTACTTGTGGACAT 5881
QY 5894 CTTGCGGGCTATGCGGACAGGGGTAGCGGCGCACTCGTGGCTTTAAGGTCATGAGCGG 5953
DB 5882 TCTGCGGGCTATGCGGCGGGGTGGCTGGCGCACTCGTGGCTTTAAGGTCATGAGCGG 5941
QY 5954 CGAGTGCCTCCACCGAGGACCTGGTCAACTTACTCTCCGCCATCTCTCTCTGGTGC 6013
DB 5942 CGAGATGCCCTCCACTGAGGATCTGGTTAATTTACTCTCTGCCATCTCTCTCTGGCGC 6001
QY 6014 CTTGGTCTGCGGGTCTGTGTCGCGAGCAATATCTGCTGCGCACGTTGGGCGCGGAGAGGG 6073
DB 6002 CTTGGTCTGCGGGTCTGTGTCGCGAGCAATATCTGCTGCGCACGTTGGGCGCGGAGAGGG 6061
QY 6074 GGCTGTGAGTGGATGAACCGGCTGTAGCGTTCGCTTCGCGGGTAAACAGCTCTCCCC 6133
DB 6062 GGCTGTGAGTGGATGAACCGGCTGTAGCGTTCGCTTCGCGGGTAAACAGCTCTCCCC 6121
QY 6134 TACGACTATGTGCTGAGAGCGACGCTGACGACGCTGCTCACTCAGATCTCTCTAGGCT 6193
DB 6122 CACGCACTATGTGCGCGAGAGCGCGCGCGGTGTGTACTCAGATCTCTCTCAGGCT 6181
QY 6194 TACCATCATCTCAACTGTGTGAAGCGGCTCCACAGTGGATTAATGAGGACTGCTCTACGCG 6253
DB 6182 TACCATCACTCAGTTGTGTGAAGAGGCTTCATCAGTGGATTAATGAGGACTGCTCTCAGCGC 6241
QY 6254 ATGCTCGGCTCGTGGCTTAAGGATGTTTGGATTTGGATATGCAAGGTTGACTGACTT 6313
DB 6242 TTGTTCCGGCTCGTGGCTTAAAGGATGTTTGGGACTGGATATGCAAGGTTGACTGACTT 6301
QY 6314 CAAGACCTGGCTCCAGTCCAACTCTCGCGCGGTTACCGGGAGTCCCTTTCTCTCATG 6373
DB 6302 CAAGACTTGGCTCCAGTCCAACTCTCGCGCGGTTACCGGGACTCTCTCTCTCTCATG 6361
QY 6374 CCAACGCGGGTACAAGGAGTCTGCGCGGGGAGCGGCATCATGCAAAACACCTGCGCCATG 6433
DB 6362 CCAACGCGGGTACAAGGAGTCTGCGCGGGGAGTGGCATCATGCAAAACACCTGCGCCATG 6421
QY 6434 CCGAGCACAGATCCCGGACATGTCAAAATTTGGTCTCATGAGGATTTGTTGGGCCAAAAC 6493
DB 6422 TGGAGCACAGATCCCGGACATGTCAAAATTTGGTCTCATGAGGATTTGTTGGGCCAAAAC 6481
QY 6494 CTGAGCAACACGTTGGCACGGAACGTTCCCATCAACGCATACACACGCGGACCTTGAC 6553
DB 6482 CTGAGCAACACGTTGGCATGGAACATTTCCCATCAACGCATACACACGCGGCCCTGAC 6541
QY 6554 ACCCTCCCGCGGCCAACTATTTCCAGGGCGTATGCGGGGTGGTGTCTGAGGAGTACGT 6613
DB 6542 GCGCTCCCGCAGCGCGAACTATTTCCAGGGCGTGTGGCGGGTGGTGTCTGAGGAGTACGT 6601
QY 6614 GAGGTTACGCTGTGGGGATTTCCCACTAGTGAAGGCGCATGACCACTGACCAACGTAAC 6673
DB 6602 GAGGTTACGCGGGTGGGGATTTCCCACTAGTGAAGGCGCATGACCACTGACCAACGTAAC 6661
QY 6674 GTGCCCATGCCAGGTTCCCGGCCCGGAATTTTTCACGAGGTTGGATGGAGTTCGGTTCGA 6733
DB 6662 ATGCCCATGCCAGGTTCCAGGCCCTGAATTTTTCACGAGGTTGGATGGAGTTCGGTTCGA 6721
QY 6734 CAGGTAAGCTCCGGCGTGCAAACTCTTCTTACGAGGAGGAGTCACTGCTTCAGGTCGGGCT 6793
DB 6722 CAGGTAAGCTCCAGTGTGCAAACTCTTCTTACGAGGAGGAGTCACTGCTTCCAGGTCGGGCT 6781
QY 6794 CAACCAATACCTTGTTCGGGTGCGAGTCCCATGCGAGCCCGAACCGGATTAACAGTGTCT 6853
DB 6782 CAACAGTACCTTGTTCGGGTGCGAGTCCCATGAGCCCGAACCGGATTAACAGTGTCT 6841
QY 6854 TACTTCCATGTCTCACCGATCCCTCCCACTTACAGAGAGGCGGCTTAAGCGTAGGCTGCGC 6913

Db	6842	CACTTCCATGCTCACCGACCCCTCTCATATATACAGCAGACGCGCAAGCTAGGCTGCG	6901
Qy	6914	TGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTGTCTCTGCGCTCTCTT	6973
Db	6902	CAGGGGTCTCCCCCTCTTGGCAGCTCTTACGTAGCCAGTGTCTGCGCTCTCTT	6961
Qy	6974	GAAGCGACATGACATACCCACATGATCCCGGACGCTGACCTCATCGAGGCCAACCT	7033
Db	6962	GAAGCGACATGATACCTACCCATCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCT	7021
Qy	7034	CTTGTGGCGCAGAGATGGCGGAACATCACTCCGCTGGAGTCAAGATTAAGTAGT	7093
Db	7022	CTTGTGGCGCAGAGATGGCGGAACATCACTCCGCTGGAGTCAAGATTAAGTAGT	7081
Qy	7094	AATTTCTGAGCTCTTTCCGAAACCGCTTCAACGCGAGGGGATGAGAGGAGATATCCGTCGC	7153
Db	7082	NATCTTGGACTCTTTCCGATCCGATTCGGGCGGTGGAGGATGAGAGGAAATATCCGTCGC	7141
Qy	7154	GGCGGAGATCTTCGGAATCCAGGAAGTTCCCTCAGCGTTGCCATATGGGACGCCC	7213
Db	7142	GGCGGAGATCTTCGGAATCCAGGAAGTTCCCTCAGCGTTGCCATATGGGACGCCC	7201
Qy	7214	GGACTACATCTCCACTGCTAGAGTCTTGAAGGACCCGGACTACGTCTCCTCGGTGGT	7273
Db	7202	GGATTACACCTCTCCACTGCTAGAGTCTTGAAGGACCCGGACTACGTCTCCTCGGTGGT	7261
Qy	7274	ACAGGATGCCATTCGCACTACCAAGGCTCTCCAAATACCACTCCAGCGAAGAG	7333
Db	7262	ACAGGATGCCATTCGCACTACCAAGGCTCTCCAAATACCACTCCAGCGAAGAG	7321
Qy	7334	GACGTTGCTCCTGACGAATCCAAATGTGCTTCTGCTTGGCGAGCTCCGCCATAAGAC	7393
Db	7322	GACGTTGCTCCTGACGAATCCAAATGTGCTTCTGCTTGGCGAGCTCCGCTACTAAGAC	7381
Qy	7394	CTTGGTAGCTCCGGATTCGTCGCGGTTGATAGCGGACCGCGACCGCCCTTCTGACCT	7453
Db	7382	CTTGGTAGCTCCGGGTCGTCGCGGTTGATAGCGGACCGCGACCGCCCTTCTGACCT	7441
Qy	7454	GGCTCCGACGAGCGGTACAAAGATCCGAGTTGAGTCCGCTCTCTCCATGCCGCCCT	7513
Db	7442	GGCTCCGACGAGCGGTACAAAGATCCGAGTTGAGTCCGCTCTCTCCATGCCGCCCT	7501
Qy	7514	TGAAGGCGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTACCGTGAGTAGGA	7573
Db	7502	CGAGGAGAGCGAGGAGACCCGACCTCAGCGACGGGTCTTGGTCTACCGTGAGCGGGA	7561
Qy	7574	GGCTAGTGAGGATGCTGCTGCTCTCAATGTCTATACGTGGACAGCGGCCCTGATCAC	7633
Db	7562	AGCTGGTGAGGACGCTGCTGCTCAATGTCTCTATACATGGACAGGTGCTTGTATCAC	7621
Qy	7634	GCCATGGCTCGGAGGAAAGTAGCTGCCCATCAACCCGTTGAGCAACTCTTTGCTGCG	7693
Db	7622	GCCATGGCTCGGAGGAGAGCAAGTTGCCCATCAATCCGTTGAGCAACTCTTTGCTGCG	7681
Qy	7694	TCACCAACAATGCTGTACGCCACAAATCCCGCAGCGCAAGCTCCGCGCAGAGAAGGT	7753
Db	7682	TCACCAACAATGCTGTACTCCAAACAATCTCGCAGCGCAAGTCTCGCGCAGAGAAGGT	7741
Qy	7754	CACCTTTGACAGATTGCAAGTCTGGATGATCATTTACGGGACGCTACTCAAGGAGATGAA	7813
Db	7742	CACCTTTGACAGATTGCAAGTCTGGACGACCACTACCGGACGCTCTCAAGGAGATGAA	7801
Qy	7814	GGGAGGCGTCCACAGTTAAGGCTAGCTCTATCTATAGAGGAGGCTCCAGCTGAC	7873
Db	7802	GGGAGGCGTCCACAGTTAAGGCTAGCTCTATCTATAGAGGAGGCTCCAGCTGAC	7861
Qy	7874	GCCCCCATTTGGCCCAATCCAAATTTGGCTATGGGCAAGGACGCTCCGGAACCTATC	7933
Db	7862	GCCCCCATTTGGCCCAATCCAAATTTGGCTATGGGCGAAGGACGCTCCGAGGCTATC	7921
Qy	7934	CAGCAGGCGCTTAACCAATCCGCTCCGCTGGGAGGACTTGTGTGGAAGACACTGAAAC	7993

Db	7922	CAGCAGGCGCTCAACCAATCCGCTCCGCTGGGAGGACTTGTGGAAGACACTGAAAC	7981
Qy	7994	ACCAATTGACACCAATCATGGCAAAAAGTAGGTTTTCTGGTCCACACAGAGAGG	8053
Db	7982	ACCAATTGATACCAATCATGGCAAAAAGTAGGTTTTCTGGTCCACACAGAGAGG	8041
Qy	8054	AGSCCGCAGCAGCTCGCTTATCGTATCCAGACCTGGGAGTTCTGTATGCGAGAA	8113
Db	8042	AGSCCGCAGCAGCTCGCTTATCGTATCCAGACCTGGGAGTTCTGTATGCGAGAA	8101
Qy	8114	GATGGCCCTTTACGAGTGTCTCCACCTCTCTCAGGCGGTGATGGGTCTCTATACGG	8173
Db	8102	GATGGCCCTTTACGAGTGTCTCCACCTCTCTCAGGCGGTGATGGGTCTCTATACGG	8161
Qy	8174	ATTTCAATCTCCCAACAGCGGCTCGATTCCTGGTGAATACCTGGAATCAAGAA	8233
Db	8162	ATTTCAATCTCTCTCGGCGGCTCGATTCCTGGTGAATACCTGGAATCAAGAA	8221
Qy	8234	ATGCCCTATGGGCTTCTCATATGACACCGCTGTTTTGACTCAACGCTCACTGAGAGTGA	8293
Db	8222	ATGCCCTATGGGCTTCTCATATGACACCGCTGTTTTGACTCAACGCTCACTGAGATGA	8281
Qy	8294	CATTCTGTGTGAGGAGTCAATTTTACCAATGTTGTGACTTTGGCCCCCGAGCCACAGGC	8353
Db	8282	CATTCTGTGTGAGGAGTCAATTTTACCAATGTTGTGACTTTGGCCCCCGAGCCACAGGC	8341
Qy	8354	CATAAGGTCTCTCAGAGCGGCTTTCATCTCGGGGTCTCTGACTCAATCACTCAAAAGGCA	8413
Db	8342	CATAAGGTCTCTCAGAGCGGCTTTCATCTCGGGGTCTCTGACTCAATTAATTCGAAGGCA	8401
Qy	8414	GAATCGGCTTATCGCGGTGCGCGCAAGTGGCTGCTGACGACTAGCTGCGGTAATAC	8473
Db	8402	GAATCGGCTTATCGCGGTGCGCGCAAGTGGCTGCTGACGACTAGCTGCGGTAATAC	8461
Qy	8474	CCTCACATGTTACTTCAAGGCCACTGCGAGCTCTGCGAGCTGCAAAAGCTCCAGGACTGCAC	8533
Db	8462	CCTCACATGTTACTTCAAGGCCACTGCGAGCTCTGCGAGCTGCAAAAGCTCCAGGACTGCAC	8521
Qy	8534	GATGCTCGTGAACCGGAGACGCTTGTGTTATCTGTGAAGCGCGGGAACCCAGGAGGA	8593
Db	8522	GATGCTCGTGAACCGGAGACGCTTGTGTTATCTGTGAAGCGCGGGAACCCAGGAGGA	8581
Qy	8594	TGCGGCGGCTTACGAGGCTTTCAGGAGGCTATGACTAGGTTATTCGCGCCCCCGGGGA	8653
Db	8582	TGCGGCGGCTTACGAGGCTTTCAGGAGGCTTTCAGGAGGCTATTCGCGCCCCCGGGGA	8641
Qy	8654	TCCGCCCCCAACAGATACGACCTTGGAGCTGATAACATCATGTTCTTCCATGTTGTCAGT	8713
Db	8642	TCCGCCCCCAACAGATACGACCTTGGAGCTGATAACATCATGTTCTTCCATGTTGTCAGT	8701
Qy	8714	CGCGCAGATGCTCATCTGGCAAAAGGTTATCTACTCCTCACCGTGACCCACCCGCCCT	8773
Db	8702	CGCGCAGATGCTCATCTGGCAAAAGGTTATCTACTCCTCACCGTGACCCACCCGCCCT	8761
Qy	8774	TGCA CGGCTGCTGGGAGACAGCTAGACACATCTCCAAATCAACTCTTTGGCTAGGCAATAT	8833
Db	8762	CGCACGGCTGCTGGGAGACAGTTAGACACATCTCCAGTCAACTCTTGGCTAGGCAATAT	8821
Qy	8834	CATCATGATGTCGCCACCCCTTATGGCAAGGATGTTCTGATGACTCACTTTTCTCCAT	8893
Db	8822	CATCATGATGTCGCCACCCCTTATGGCAAGGATGTTCTGATGACTCACTTTTCTCCAT	8881
Qy	8894	CCTTCTAGCTCAAGAGCACTTGAAAAAGCCCTGGATTTGTGAGATCTACGGGCTTGCTA	8953
Db	8882	CCTTCTAGCTCAAGAGCACTTGAAAAAGCCCTGGATTTGTGAGATCTACGGGCTTGCTA	8941
Qy	8954	CTCCATTTGAGCCACTTGACCTTACCTCAGATCAITTTGAACGACTCCATGGTCTTAGGGCAT	9013
Db	8942	CTCCATTTGAGCCACTTGACCTTACCTCAGATCAITTTGAACGACTCCATGGTCTTAGGGCAT	9001
Qy	9014	TACACTCCAGTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCTGCTCAGGAACT	9073
Db	9002	TTCACTCCACAGTTACTCTCCAGGTGAGATCAATAGGTTGGCTTCTGCTCAGGAACT	9061

1394 CCACTGGGAGTCTCTGGCGGCTTCCCTACTATTCATGGTAGGAACTGGGCTAAGGT 1453
1382 CCACTGGGAGTCTCTAGCGGCTTCCCTACTATTCATGGTAGGAACTGGGCTAAGGT 1441
1454 TCTGATGTGGCGCTACTCTTTGGCGGCTTGACGGGAGACCCACACACGGGAGGCT 1513
1442 CTTGATGTGATGCTACTCTTTGGCGTGTGACGGGACACCCACGTGACAGGGGAAG 1501
1514 GGCGGCGCACACCACTCCGGGTTCAGTCCCTTTCTCATCTGGGCGCTCTCAGAAAT 1573
1502 GGTAGCTCCAGCACCCAGAGCCTGTGTCTCTGCTCTCACAAGGCCCATCTCAGAAAT 1561
1574 CCACTGTGTGAATPACCAACGGCAGCTGGCACAATCAACAGGACTGCCCTAAATTCGAATGA 1633
1562 CCAACTCGTGAAACACCAACGGCAGCTGGCACAATCAACAGGACTGCCCTGAAATTCGAATGA 1621
1634 CTCCTCTCAAACTCGGTTCTTTGGCGGCTGTTTACGACACAAAGTTCAACTCGTCCGG 1693
1622 CTCCTCTCAAACTCGGTTCTTTGGCGGCTGTTTACGACACAAAGTTCAACTCGTCCGG 1681
1694 GTCCCGGAGCGATGCGCAGCTGCGGCCCATTTGACTGCTCGCCAGGGGTGGGCCC 1753
1682 GTCCCGAGCGATGCTAGCTGCGGCCCATTCGATGATGCTCAGGGGTGGGTCC 1741
1754 CATCACTATACTAAGCTTAACAGCTCGGATCAGAGGCTTATTTGCTGGCATTAACGGCC 1813
1742 CATCACTCATGATATGCTCAGAGCTCGGACAGAGGCCATATTTGCTGGCACTACGGGCC 1801
1814 TCACCGCTGTGTGTCTACCGCGTCGAGGTGTGTGTCAGTGTATTTGTTTCACCCC 1873
1802 TCACCGTGTGGGATCGTGTGCTGCGAGGTGTGTGTCAGTGTATTTGTTTCACTCC 1861
1874 AGCCCTGTGTGTGGGAGCACCGATGTTCCGCTGCTCCTACGATAGCTGGGGGA 1933
1862 GAGCCCTGTGTGTGGGAGCACCGATGTTCCGCGCTCCTACGATAGCTGGGGGA 1921
1934 GAATGACACAGCTGATGCTCTCAACACACGCGTCCGCAACAGGCACTGGTTCCG 1993
1922 GAATGACACAGCTGATGCTCTCAACACACGCGGCTCCAGGCACTGGTTGG 1981
1994 CTGTACATGATGAATAGTACTGGGTTCACTAAGACGTGGGAGGTCCTCCCGTGTAAACAT 2053
1982 GTGACGTGTGATGAACAGCACTGGTTCAACAGACGTGGGGGCTCCGTGCAACAT 2041
2054 CGGGGGGTGCGGTAAACCGCACTTTGATCTGCCCCACGAGCTGCTTCGGAAGCAACCCGA 2113
2042 CGGGGGGTGCGGAACCAACACTTTGGTCTGCCCCACGAGTTGCTTCGGAAGCAACCCGA 2101
2114 GGCTACTTACACAAATGTGTGCTCGGGGCTGCTGTGACACCTAGGTGCTTAGACTA 2173
2102 GGCCACTTACACAAATGTGTGCTCGGGGCTGCTGTGACACCTAGGTGCTTAGACTA 2161
2174 CCGATACAGGCTTTGGCACTACCTGCACTCTCAATTTTCCATCTTTAAGGTAGGAT 2233
2162 CCGATACAGGCTCTGGCACTACCTGCACTCTTTAAGGTAGGAT 2221
2234 GTATGTGGGGGTGGGAGCAGGCTCAATGCGCGCATGCAATTTGCACTCGAGGAGCG 2293
2222 GTATGTGGGGGTGGGAGCAGGCTCAATGCTGATGCAATTTGCACTCGAGGAGCG 2281
2294 CTGTAACTTGGAGCAGGATAGGTGAGAACTCAGCCCGCTGCTGTGTCTACAAACA 2353
2282 CTGTGACTTGGAGCAGGATAGGTGAGAACTCAGCCCGCTGCTGTGTCTACAAACA 2341
2354 GTGGCAGATATGCGCTGTGCTTTCACACCTACCGGCTTTATCCACTGGTTGATCCA 2413
2342 GTGGCAGATATGCGCTGTGCTTTCACACCTACCGGCTTGTCCACTGGCTTGATCCA 2401
2414 TCTCCATCAGAACATCGTGGAGCTGCAATCTGTACGCTGTAGGCTCAGGCTTGTCTC 2473
2402 TCTTACCGGAACATCGTGGAGCTGCAATCTGTACGCTGTAGGCTCAGGCTTGTCTC 2461

2474 CTTTGCATCAAAATGGAGTACATCTGTGTCTTTTCTCTCTCTGGCAGACGCGCGCT 2533
2462 CTTTGCATCAAAATGGAGTATATCTGTGTCTTTTCTCTCTCTGGCAGACGCGCGCT 2521
2534 GTGTGCTCTTGTGTGATGATGCTGATAGCCAGGTGAGGCGCTTAGAGAACTT 2593
2522 CTGTGCTCTTGTGTGATGATGCTGATAGCCAGGTGAGGCGCTTAGAGAACTT 2581
2594 GGTGTCTCTCAATGCGGCTCGTGGCGGAGGATGATTTCTCTCTCTCTCTCTCTGT 2653
2582 GGTGTCTCTCAATGCGGCTCGTGGCGGAGCGCATGCGCTTCTCTCTCTCTCTGT 2641
2654 CTTCTGCGCGCTGTGTACATTAAGGCGAGGTGCGCTCTCTGGGCGGCGTATGCTTTTA 2713
2642 CTTCTGCGCGCTGTGTACATTAAGGCGAGGTGCGCTCTCTGGGCGGCGTATGCTCTTA 2701
2714 TGGCGTATGCGCGCT 2773
2702 TGGCGTATGCGCGCT 2761
2774 CCGGAGATGCGTGTGATCGTGGGCGGTGCGGTCTCTGTAGTCTGTGATTTCTTGACCTT 2833
2762 CCGAGAGATGCGTGTGATCGTGGGCGGTGCGGTCTCTGTAGTCTGTGATTTCTTGACCTT 2821
2834 GTCCACATCTACAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2893
2822 GTCCACATCTATTAAGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2881
2894 CAGAGCGGCGGCACATGCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 2953
2882 CAGAGCGGCGGCACATGCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 2941
2954 TGCCATCATCT 3013
2942 TGCCATCATCT 3001
3014 CTTGTCTGCGCATCTGCGCGGCTCATGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3073
3002 CTTGTCTGCGCATCTGCGCGGCTCATGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3061
3074 CTTGTCTGCGCATCTGCGCGGCTCATGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3133
3062 CTTGTCTGCGCATCTGCGCGGCTCATGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3121
3134 TCATTATGTCCAAATGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3193
3122 CCATATGTCCAAATGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3181
3194 CCATCTTACCCCACTGCGGCACTGGGCGCCACGCGGCGCTACGAGACCTTGGCGGT 3253
3182 CCATCTTACCCCACTGCGGCACTGGGCGCCACGCGGCGCTACGAGACCTTGGCGGT 3241
3254 AGAGCCCGTCT 3313
3242 AGAGCCCGTCT 3301
3314 TGCGTGTGGGACATCATCTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3373
3302 GGCGTGTGGGACATCATCT 3361
3374 TTTGGGACCGGCTCATAGTCTCGAAGGCGAAGGCTGCGGCTCTCTCTCTCTCTCTCTCTCT 3433
3362 TCTAGGACCGGCTCATAGTCTCGAAGGCGAAGGCTGCGGCTCTCTCTCTCTCTCTCTCTCT 3421
3434 CTACTCTCCCAACAAACGGGCGCTACTTGGTGTGATCATCTAGCTCTACAGGCGGGA 3493
3422 CTATTTCCCAACAAACGGGCGCTACTTGGTGTGATCATCTAGCTCTACAGGCGGGA 3481
3494 CAAGAACAGGTGCAAGGGAGGTTCAGTGGTGTCTACCGCAACAACTCTTCTCTCTCTCT 3553
3482 CAAGAACAGGTGCAAGGGAGGTTCAGTGGTGTCTCTCCACCGCAACGCAATCTTCTCTCT 3541
3554 GACCTGCAATCAACGGCGTGTCTGATGTCTACCATGCGGCTGCTCTGCAAGACCTTAGC 3613

[illegible]

Db	4622	CGTCGTTGTCGTGGCAACACAGACGCTCTAATGACGGGTTTTACCGGCGACCTTTGACTCACTG	4681
Qy	4694	GATCGACTGCAATACATGTGTCAACCCAGACAGTCGACTTCAGCTTGGATCCCACTTCAC	4753
Db	4682	GATCGACTGCAACACATGTGTCAACCCAGACAGTCGATTTTCAGCTTGGATCCCACTTCAC	4741
Qy	4754	CATTGAGACAGACGACCGTGCCCAAGACGCGGTGTGCGCTTCGGAACGGGAGGTAGAAC	4813
Db	4742	CATTGAGACACAACGCTGCCCCAAGACGCGGTGTGCGGTGCGCAGCGCGGAGGTAGGAC	4801
Qy	4814	TGGCAGGGGTAGGATGGGCATCTACAGGTTTGTGACTCCAGGAGAGACGGGCCCTCGGGCAT	4873
Db	4802	TGGCAGGGGCAGGATGGGCATCTACAGGTTTGTGACTCCAGGAGAGACGGGCCCTCAGGCAT	4861
Qy	4874	GTTTCGATTTCTTCGGTCCCTGTGTGAGTGCTATGACGCGGCTGTCTTGGTATGAGCTCAC	4933
Db	4862	GTTTCGACTCTCTCGTCTGTGTGAGTGCTATGACGAGGCTGCGCTTGGTATGAGCTCAC	4921
Qy	4934	GCCCGCTGAGACCTCGGTTAGGTTGGGGCTTACTTAATACACAGGGTTGCCGCTGTG	4993
Db	4922	GCCCGCTGAGACCTCGGTTAGGTTGGGGCTTACTTAATACACAGGGTTGCCGCTGTG	4981
Qy	4994	CCAGGACCATCTGGAGTCTTGGGAGAGCGCTTTACAGGGCTCACCCATAGATGCCCA	5053
Db	4982	CCAGGACCATCTAGAGTCTTGGGAGAGCGCTTTACAGGGCTCACCCATAGATGCCCA	5041
Qy	5054	CTTCCTGTGCCAGACTAAACAGGCAGGAGACAACTTTCTTACTCTGGTGGCATATCAAGC	5113
Db	5042	CTTCCTGTGCCAGACCAACAGGCAGGAGAGAACCTTCCCTACTCTGGTAGCATACCAAGC	5101
Qy	5114	TACAGTGTGGGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAAATGGAAGTGCT	5173
Db	5102	CACAGTGTGGGCCAGGGCTCAGGCTCCACCTCCATCGTGGGACCAAAATGGAAGTGCT	5161
Qy	5174	CATPAGGCTGAACCTTACATGACGCGGGCAACACCCCTGCTGTATAGGCTAGGAGCGGT	5233
Db	5162	CATPAGGCTGAAGGCCACACTGCTATGGGCCAACGCCCCCTGCTGTACAGGCTAGGAGCGGT	5221
Qy	5234	CCAAATGAGGTCACTCTCACACCCCACTAACTTAAATACATCATGCGATGCATGTCGGC	5293
Db	5222	TCAAATGAGGTCACTCTCACACCCCACTAACTTAAATACATCATGCGATGCATGTCGGC	5281
Qy	5294	TGACCTGGAGGTCTGCTACTAGCACCTGGGTGCTGGTAGGGCGAGTCTCTGCAGCTTTGGC	5353
Db	5282	TGACCTGGAGGTCTGCTACTAGCACCTGGGTGCTGGTAGGGCGAGTCTCTGCAGCTTTGGC	5341
Qy	5354	CGCATCTGCTGACGACAGCGAGTGTGTCATTGTGGGAGGATCATCTTGTCCGGGAA	5413
Db	5342	CGCGTACTGCTGACGACAGCGAGTGTGTCATTGTGGGAGGATCATCTTGTCCGGGAG	5401
Qy	5414	GCCAGCTGTGTTCCCGACAGGGAGTCTCTACAGGAGTTCGATGAGATGGAAGAGTG	5473
Db	5402	GCCAGCTGTATTCCCGACAGGGAGTCTCTACAGGAGTTCGATGAGATGGAAGAGTG	5461
Qy	5474	TGCCTCACACTTCTTTACATCGAGCAGGGAATGCAGCTCGCCGAGCAATTCGAAGCAAAA	5533
Db	5462	TGCTTCACACTTCTTACATCGAGCAGGGAATGCAGCTCGCCGAGCAATTCGAAGCAAAA	5521
Qy	5534	GGCGCTCGGGTTGTGCAAAACGGCCACCAAGCAACGGAGGCTGCTCTCCCGTGGTGA	5593
Db	5522	GGCGCTCGGATTGTGCAAAACAGCCACCAAGCAACGGAGGCTGCTCTCCCGTGGTGA	5581
Qy	5594	GTCNAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCAATGTGGAAATTCATCAGGG	5653
Db	5582	GTCNAAGTGGGAGCCCTTGAGGCTTCTGGGCGAAGCAATGTGGAAATTCATCAGGG	5641
Qy	5654	AATACAGTACCTAGCAGGCTTATCCACTCTGCTCTGGAAAACCCCGCATAGCATCATTTGAT	5713
Db	5642	GATACAGTACTTGGCAGGCCTATCCACTCTGCTCTGGAAAACCCCGCATAGCATCATTTGAT	5701
Qy	5714	GGCATTTACAGCTTCTATCACTAGCCCGCTCAACCAACCAAAACACCTCTCTGTTTAAAT	5773
Db	5702	GGCTTTTACAGCTTCTATCAACAGCCCGCTCAACCAACCAAAATACCTCTCTGTTTAAAT	5761

QY	5774	CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGTGCCTCAGCTTTCTGTGG	5833
DB	5762	CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGTGCCTCAGCTTTCTGTGG	5821
QY	5834	CGCCGGATCGCGGAGCGGTGGTGGCAGCATAGGCTTGGGAAGTGTCTGTGACAT	5893
DB	5822	CGCCGGATCGCGGAGCGGTGGTGGCAGCATAGGCTTGGGAAGTGTCTGTGACAT	5881
QY	5894	CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGTGCCTCAGCTTTCTGTGG	5953
DB	5882	CTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGTGCCTCAGCTTTCTGTGG	5941
QY	5954	CGAGGTGCTCTCCAGGAGCGGTGGTGGCAGCATAGGCTTGGGAAGTGTCTGTGAC	6013
DB	5942	CGAGGTGCTCTCCAGGAGCGGTGGTGGCAGCATAGGCTTGGGAAGTGTCTGTGAC	6001
QY	6014	CTTGGGTGCTGGGGTGGTGGCAGCAATCTGCGTGGGAGGAGGAGG	6073
DB	6002	CTTGGGTGCTGGGGTGGTGGCAGCAATCTGCGTGGGAGGAGGAGG	6061
QY	6074	GGCTGTCAGTGGATGAACCGGTGTAGCGTTTCGCTTGGGGGTAAACACGCTCTCC	6133
DB	6062	GGCTGTCAGTGGATGAACCGGTGTAGCGTTTCGCTTGGGGGTAAACACGCTCTCC	6121
QY	6134	TAGCACTATGCTGCTGAGAGCGAGCTGCGAGCACTGTCTCACTCAGATCTCTTAGCCT	6193
DB	6122	CACGCACTATGCTGCTGAGAGCGAGCTGCGAGCACTGTCTCACTCAGATCTCTTAGCCT	6181
QY	6194	TACCATCACTCACTGCTGAAGCGGCTCCACAGTGGATTAATGAGAGCTGCTTAGCCT	6253
DB	6182	TACCATCACTCACTGCTGAAGCGGCTCCACAGTGGATTAATGAGAGCTGCTTAGCCT	6241
QY	6254	ATGCTCGGCTCGTGGCTTAAGGATGTTGGATGTTGATGATGACCGTGTGAGTACTT	6313
DB	6242	TTGTTTCGGCTCGTGGCTTAAGGATGTTGGATGTTGATGATGACCGTGTGAGTACTT	6301
QY	6314	CAAGACCTGCTCGAGTCCAACTCTCGCGCGGTACCGGAGTCCCTTTCTGTGATG	6373
DB	6302	CAAGACCTGCTCGAGTCCAACTCTCGCGCGGTACCGGAGTCCCTTTCTGTGATG	6361
QY	6374	CCAAAGCGGGTACAAAGGAGTCTGGCGGGGAGCGGATCATGCAACACCTGCGCATG	6433
DB	6362	CCAAAGCGGGTACAAAGGAGTCTGGCGGGGAGCGGATCATGCAACACCTGCGCATG	6421
QY	6434	CGGAGCACAGATCGCGGACATGTCAAAACGCTTCATAGGATCGTAGGCGCTAGAAC	6493
DB	6422	TGGAGCACAGATCGCGGACATGTCAAAACGCTTCATAGGATCGTAGGCGCTAGAAC	6481
QY	6494	CTGAGCAACACGCTGGACGGAACGCTTCCCATCAACGATACACCGGAGCTTGGCAC	6553
DB	6482	CTGAGCAACACGCTGGACGGAACGCTTCCCATCAACGATACACCGGAGCTTGGCAC	6541
QY	6554	ACCTTCCCGCGGCGCAACTATTCAGGGCGCTATGGCGGTGGCTGTGAGGATAGT	6613
DB	6542	GCCCTCCCGCGGCGCAACTATTCAGGGCGCTATGGCGGTGGCTGTGAGGATAGT	6601
QY	6614	GGAGGTACGCTGTGGGGATTTCCACTACGTGACGGGATGACCACTGACACGTAATA	6673
DB	6602	GGAGGTACGCTGTGGGGATTTCCACTACGTGACGGGATGACCACTGACACGTAATA	6661
QY	6674	GTGCCCACGAGGTTCGGGCGGCGGATTTCTTACGGAGGTGGATGGAGTGGGTTGCA	6733
DB	6662	ATGCCCACGAGGTTCGAGCGGCGGATTTCTTACGGAGGTGGATGGAGTGGGTTGCA	6721
QY	6734	CAGGTACGCTCGCGGCGTCAAACTCTTCTACGGGAGGACGCTCAGCTTCCAGGTCCGGCT	6793
DB	6722	CAGGTATGCTTCAGTGTGCAACCTCTCTACGAGGAGGTCTGTATTCAGGTCCGGCT	6781
QY	6794	CAACCAATATCTGGTGGGTGCGAGCTCCCATGCGAGCCGGAACCGGAGCTAACAGTGT	6853
DB	6782	CAACCAATATCTGGTGGGTGCGAGCTCCCATGCGAGCCGGAACCGGAGTGTGGAGTGT	6841

QY	6854	TACTTCCATGCTACCGATCCCTCCACATTTACGACAGACGGCTAAGCTAGGCTGGC	6913
DB	6842	CACCTTCCATGCTACCGATCCCTCCACATTTACGACAGACGGCTAAGCTAGGCTGGC	6901
QY	6914	TAGAGGTCTCCCGCTCTTTTAGCCAGCTCATCAGCTAGCCAGTGTGTGCGCTCTTT	6973
DB	6902	CAGGGGTCTCCCGCTCTTTTAGCCAGCTCATCAGCTAGCCAGTGTGTGCGCTCTTT	6961
QY	6974	GAAAGGACATGACATACCCACATGACTCCCGGACGCTGACCTCATCGAGGCCAACCT	7033
DB	6962	GAAAGGACATGACATACCCACATGACTCCCGGACGCTGACCTCATCGAGGCCAACCT	7021
QY	7034	CTTGTGCGCGCAGGAGATGGCGGAAACATCCTCCTCGCTGGAGTCAGAGAAATAGGTAGT	7093
DB	7022	CTTGTGCGCGCAGGAGATGGCGGAAACATCCTCCTCGCTGGAGTCAGAGAAATAGGTAGT	7081
QY	7094	AATTTGAGACTCTTTTCGAAACCGCTTCCAGCGGAGGGATGAGAGGAGATATCCGTGCG	7153
DB	7082	AATTTGAGACTCTTTTCGAAACCGCTTCCAGCGGAGGGATGAGAGGAGATATCCGTGCG	7141
QY	7154	GGCGGAGATCTCGCGGAAATCCAGGAAGTTCCCTCAGCGTTGCCCATATGGGACGCCC	7213
DB	7142	GGCGGAGATCTCGCGGAAATCCAGGAAGTTCCCTCAGCGTTGCCCATATGGGACGCCC	7201
QY	7214	GGACTACATCTCTCCACTGCTAGAGTCTTGGAAAGGACCCGAGACTACGTCTCCCTCGGTGGT	7273
DB	7202	GGATTAACACCTCTCCACTGCTAGAGTCTTGGAAAGGACCCGAGACTACGTCTCCCTCGGTGGT	7261
QY	7274	ACACGAGTGGCAATGCTACCAAGGCTCTCCAATACCACTCCACGGAAGAAAGAG	7333
DB	7262	ACACGAGTGGCAATGCTACCAAGGCTCTCCAATACCACTCCACGGAAGAAAGAG	7321
QY	7334	GACGGTGTCTGACAGAAATCCAAATGTCTTCTGCTTGGCGGAGCTCGCCACTAGAC	7393
DB	7322	GACGGTGTCTGACAGAAATCCAAATGTCTTCTGCTTGGCGGAGCTCGCCACTAGAC	7381
QY	7394	CTTTCGTAGCTCGGATCGCTGCGCGTGTAGCGGACGCGGACCGCCCTTCTCGTACCT	7453
DB	7382	CTTTCGTAGCTCGGATCGCTGCGCGTGTAGCGGACGCGGACCGCCCTTCTCGTACCT	7441
QY	7454	GGCTTCCGACGAGGTCGACAAAGGATCCGAGCTTGGAGTCTCTCTCCATGCCCCCT	7513
DB	7442	GGCTTCCGACGAGGTCGACAAAGGATCCGAGCTTGGAGTCTCTCTCCATGCCCCCT	7501
QY	7514	TGAAGGGAGCGCGGGGACCCCGATCTCAGCGAGGCTTTGGTCTACCGTGTAGGAGA	7573
DB	7502	CGAGGGAGAGCGAGGGGACCCCGATCTCAGCGAGGCTTTGGTCTACCGTGTAGGAGA	7561
QY	7574	GGCTAGTGGAGTGTCTGCTGCTCAATGCTTATAGCGGACGCGGCTTGTATCAC	7633
DB	7562	AGCTGTGAGGAGTGTCTGCTGCTCAATGCTTATAGCGGACGCGGCTTGTATCAC	7621
QY	7634	GCCATGCTGCGGAGGAAAGTAAAGCTGCCATCAACCGCTTGGAGCAACTCTTTGCTGCG	7693
DB	7622	GCCATGCTGCGGAGGAGGAAAGTAAAGCTGCCATCAACCGCTTGGAGCAACTCTTTGCTGCG	7681
QY	7694	TCACCAAAACATGCTCTACGCGCAACATCCCGAGCGCAAGCTTCCGCGAGAGAAAGGT	7753
DB	7682	TCACCAAAACATGCTCTACGCGCAACATCCCGAGCGCAAGCTTCCGCGAGAGAAAGGT	7741
QY	7754	CACCTTTGACAGATGCAAGTCTCGGATGATCATTCAGGAGCGCTACTCAGAGAGTAA	7813
DB	7742	CACCTTTGACAGATGCAAGTCTCGGATGATCATTCAGGAGCGCTACTCAGAGAGTAA	7801
QY	7814	GGCGAGGCTGCGAGGCTTAAAGCTTCTATATAGAGGAGGCTTCAAGCTGAC	7873
DB	7802	GGCGAGGCTGCGAGGCTTAAAGCTTCTATATAGAGGAGGCTTCAAGCTGAC	7861
QY	7874	GCCCCCAATTCGGCGCAAAATCCAAATTTGGCTATGGGCGAAAGGACGCTCCGGAACCTATC	7933
DB	7862	GCCCCCAATTCGGCGCAAAATCCAAATTTGGCTATGGGCGAAAGGACGCTCCGGAACCTATC	7921
QY	7934	CAGCAGGCGCTTAAACCAATCTCCGCTCGTGTGGGAGGACTTGTGGAGAGACACTGAAAC	7993

[illegible]

Db	9002	TTCACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTTCATGCTCAGGAAC	9061
Qy	9074	TGGGGTACCACCCCTTCGCAACCTGGAGACATCGGGCCAGAAAGTGTCCGGCTAAGCTTACT	9133
Db	9062	TGGGGTACCAGCTTTTCGAGTCTGGAGACATCGGGCCAGAAAGTGTCCGGCTAAGCTTACT	9121
Qy	9134	GTCCCAAGGGGGGAGGGCCGCGACATTTGTGGCAGATACCTCTTTAACTGGGCAGTAAGGAC	9193
Db	9122	GTCCCAAGGGGGGAGGGCTGCCACTTTGCGGCAAGTACCTCTTCAACTGGGCAGTAAGGAC	9181
Qy	9194	CAAGCTTTAAACTCACAATCCCGCGCGTCCCAAGCTGGAGCTTGTCTGGCTGGTTTCGT	9253
Db	9182	CAAGCTTTAAACTCACTTCCAATCCCGCGTCCCAAGCTGGAGCTTGTCTGGCTGGTTTCGT	9241
Qy	9254	CGCTGGTTACAGCGGGGAGACATATATACACGCCCTGTCTCGTCCCGACCCCGCTGGTT	9313
Db	9242	TGCTGGTTACAACGGGGGAGACATATATACACGCCCTGTCTCGTCCCGACCCCGCTGGTT	9301
Qy	9314	TCCGTTGTGCTACTCTTACTTTCTGTAGGGGTAGGCAATTTACTGTCCTCCCAACCGATG	9373
Db	9302	CATGTTGTGCTACTCTTACTTTCTGTAGGGGTAGGCATCTACCTGCTCTCCCAACCGGTG	9361
Qy	9374	AACGGGAGCTAACCACTCCAGGCGCTTAAGCCATTTCTCGTTTTTTTTTTT 9424	
Db	9362	AACGGGAGCTAACCACTCCAGGCGCTTAAGCCATTTCTCGTTTTTTTTTTT 9412	
RESULT 3			
US-10-467-000-2			
; Sequence 2, Application US/10467000			
; Publication No. US20040067486A1			
; GENERAL INFORMATION:			
; APPLICANT: De Francesco, Raffaele			
; APPLICANT: Migliaccio, Giovanni			
; APPLICANT: Pronessa, Giacomo			
; TITLE OF INVENTION: HEPATITIS C VIRUS REPLICONS AND REPLICON			
; FILE OF INVENTION: ENHANCED CELLS			
; FILE REFERENCE: ITR0003P			
; CURRENT APPLICATION NUMBER: US/10/467,000			
; CURRENT FILING DATE: 2003-07-21			
; PRIOR APPLICATION NUMBER: PCT/EP02/00526			
; PRIOR FILING DATE: 2002-01-16			
; PRIOR APPLICATION NUMBER: 60/263,479			
; PRIOR FILING DATE: 2001-01-23			
; NUMBER OF SEQ ID NOS: 13			
; SOFTWARE: Fast-Seq for Windows Version 4.0			
; SEQ ID NO 2			
; LENGTH: 9605			
; TYPE: DNA			
; ORGANISM: Con 1 HCV isolate amino acid			
US-10-467-000-2			

1

Db 241 |||||GCAGACTCTAGCCGAGTAGTGTGGGTCGGGAAAGGCTTGTGGTACTGGCTGTATGG 300
Qy 301 GTGCTTGCAGAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGAAATCCCTAAAC 360
Db 301 GTGCTTGCAGAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGAAATCCCTAAAC 360
Qy 361 CTCAAGAAAAACAAACGCTAACAACACCGCGCCGACAGGACGTAAGTTCCCGGGG 420
Db 361 CTCAAGAAAAACAAACGCTAACAACACCGCGCCGACAGGACGTAAGTTCCCGGGG 420
Qy 421 GTGGTCAGATCGTTGCTGGAGTTTACTGTTGCGCGGACAGGAGCCACAGTTGGGTGTC 480
Db 421 GTGGTCAGATCGTTGCTGGAGTTTACTGTTGCGCGGACAGGAGCCACAGTTGGGTGTC 480
Qy 481 GCGCGACTAGGAAGGCTTCGAGCGGTTCGCAACCTCGTGAAGGCGACAACTATCCCAA 540
Db 481 GCGCGACTAGGAAGACTTCGAGCGGTTCGCAACCTCGTGAAGGCGACAACTATCCCAA 540
Qy 541 AGGCTCGCCGACCCGAGGGCAGGGCCTGGGCTCAGCCCGGGTACCCTTGGCCCTCTATG 600
Db 541 AGGCTCGCCAGCCGAGGGTAGGGCTGGGCTCAGCCCGGGTACCCTTGGCCCTCTATG 600
Qy 601 GCAATGAGGCGCTGGGGTGGGAGGATGGCTCTCTGTACCCCGGGCTCCCGGCTAGTT 660
Db 601 GCAATGAGGCGCTGGGGTGGGAGGATGGCTCTCTGTACCCCGGGCTCCCGGCTAGTT 660
Qy 661 GGGGCCCCACGGACCCCGGCGTAGGTGCGGTAACTTGGGTAAAGTTCATCGATACCCCTA 720
Db 661 GGGGCCCCACGGACCCCGGCGTAGGTGCGGCAATTTGGGTAAAGTTCATCGATACCCCTA 720
Qy 721 CATGCGGCTTCGCGGATCTCATGGGTAACATTCGCTCGTCGGCGCCCTCAGGGGCG 780
Db 721 CGTGGCGCTTCGCGGATCTCATGGGTAACATTCGCTCGTCGGCGCCCTCAGGGGCG 780
Qy 781 CTGCAAGGCTTGGGCAACGCTGTCGGGTTCTGAGGAGCGGGTGAATGCAACAG 840
Db 781 CTGCAAGGCTTGGGCAACGCTGTCGGGTTCTGAGGAGCGGGTGAATGCAACAG 840
Qy 841 GGAATCTGCGCGGTGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 841 GGAATCTGCGCGGTGCT 900
Qy 901 TCCAGCTTCGCTTATGAAGTGGCAACGCTGTCGGGATATACCATGTCAACGACGT 960
Db 901 TCCAGCTTCGCTTATGAAGTGGCAACGCTGTCGGGATATACCATGTCAACGACGT 960
Qy 961 GCTCCAACTCAAGCAATGTGTATGAGCAGCGGACGTGATCATGCTCCCGGGTGG 1020
Db 961 GCTCCAACTCAAGCAATGTGTATGAGCAGCGGACGTGATCATGCTCCCGGGTGG 1020
Qy 1021 TGGCTGTGTTACAGAGGTAAAGCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCG 1080
Db 1021 TGGCTGTGTTACAGAGGTAAAGCTCCGCTGCTGGGTAGCGCTCACTCCACGCTCG 1080
Qy 1081 CGGCGAGAAATGCGAGGTCGCCACTACGAAATACGACGCGACGTGACGTTCGCTTG 1140
Db 1081 CGGCGAGAAATGCGAGGTCGCCACTACGAAATACGACGCGACGTGACGTTCGCTTG 1140
Qy 1141 GGAGCGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTCTCGCGATCTATTTCTCTG 1200
Db 1141 GGAGCGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTCTCGCGATCTATTTCTCTG 1200
Qy 1201 TCTCCAGCTGTTTCACTTCTCGCTCCCGGATGAGACAGTGCAGGACTGCAACTGCT 1260
Db 1201 TCGCCAGCTGTTTCACTTCTCGCTCCCGGATGAGACAGTGCAGGACTGCAACTGCT 1260
Qy 1261 CAATCTATCCCGGCACTGATCAGTCAACGCTGCTGGGATGATGATGATGATGATGATGAT 1320
Db 1261 CAATCTATCCCGGCACTGATCAGTCAACGCTGCTGGGATGATGATGATGATGATGATGAT 1320
Qy 1321 CACCTACAAAGCCCTAGTGGTGTGCGCAAGTTGCTCCCGATCCCAAGCTGCTCGTGACA 1380

Db 1321 CACCTACAGCAGCCCTAGTGGTATCGAGTTACTCCGATCCCAAGCTGCTGTGATA 1380
Qy 1381 TGGTGGCGGGGCCCACTGGGGAGTCTTGGCGGCTTGGCTACTATTTCATGGTAGGA 1440
Db 1381 TGGTGGCGGGGCCCACTGGGGAGTCTTGGCGGCTTGGCTACTATTTCATGGTAGGA 1440
Qy 1441 ACTGGGCTAAGGTTCTGATTGGGCTACTCTTTGCGGCGGTGACGGGAGACCCACA 1500
Db 1441 ACTGGGCTAAGGTTCTGATTGGGCTACTCTTTGCGGCGGTGACGGGAGACCTATG 1500
Qy 1501 CGACGGGAGGGTGGCGGCCACACACCTCCGGGTTCACTGCTCCCTTTCTCATCTGGGG 1560
Db 1501 TGACAGGGGGACGATGGCCAAAAACACCTCGGATTAAGTCCCTCTTTTCAACCCGGGT 1560
Qy 1561 CGTCTCAGAAAAATCAGCTTGTGAATACCAACGGGAGCTGGACATCAACAGACACTGCC 1620
Db 1561 CATCCAGAAAAATCCAGCTTGTAAACACCAACGGGAGCTGGACATCAACAGACACTGCC 1620
Qy 1621 TAAATTGCAATGACTCCCTCCAACTGGGTTCTTTGCGGCGTGTGTTTACGCACACAAGT 1680
Db 1621 TGAATGCAATGACTCCCTCAACACTGGGTTCTTGTGGCTGTTTCTAGCTGCACAAAT 1680
Qy 1681 TCAACTCGTCCGGGTGCCCGAGCGCATGGCCAGCTGCCGCCCATTTGACTGGTTCGCC 1740
Db 1681 TCAACTCATCTGGATGCCAGAGCGCATGGCCAGCTGCAGCCCATCGACGCGTTCGCTC 1740
Qy 1741 AGGGTGGGGCCCCATCACTATAAGCTTAACAGCTTCGGATCAGAGGCTTATTTGCT 1800
Db 1741 AGGGTGGGGCCCCATCACTATAAGCTTAACAGCTTCGGATCAGAGGCTTATTTGCT 1800
Qy 1801 GGCATTTACGGCTCGACCGTGTGTCGTACCCGCTCGCAGGTGTGTGTCAGTGT 1860
Db 1801 GGCATTTACGGCTCGACCGTGTGTCGTACCCGCTCGCAGGTGTGTGTCAGTGT 1860
Qy 1861 ATTGTTTACCCCAAGCCCTGTTGTTGGTGGGACCAACCGATCGTTCGGTGTCCCTACGT 1920
Db 1861 ACTGCTTACCCCAAGCCCTGTTGTTGGTGGGAGCAGCCGATCGTTCGGCTCCCTACGT 1920
Qy 1921 ATAGCTGGGGGAGATGAGACAGAGCTGATGTCTCTCAACAAACGCGTCCGCCACAAG 1980
Db 1921 ACAGTTGGGGGAGATGAGACAGAGCTGATGTCTCTTAAACAACGCGCGCGCGCAAG 1980
Qy 1981 GCAACTGGTTCGGCTGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Db 1981 GCAACTGGTTCGGCTGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Qy 2041 CCCGCTGTAACTCGGGGGTCCGTAACCGCACTTGAATGCTGCCCAACGACGTCTTCC 2100
Db 2041 CCCGCTGTAACTCGGGGGATCGGCAATAAACCTTGAATGCTGCCCAACGACGTCTTCC 2100
Qy 2101 GGAAGCAACCCGAGGCTTACTTACAAAAATGTGGCTCGGGGCCCTGTTGACACTAGGT 2160
Db 2101 GGAAGCAACCCGAGGCTTACTTACAAAAATGTGGCTCGGGGCCCTGTTGACACTAGGT 2160
Qy 2161 GCCTAGTAGACTACCCATACAGGCTTGGCACTACCCCTGCACTCTCAATTTTCCATCT 2220
Db 2161 GCCTAGTAGACTACCCATACAGGCTTGGCACTACCCCTGCACTCTCAATTTTCCATCT 2220
Qy 2221 TTAAGGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
Db 2221 TTAAGGTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
Qy 2281 CTGAGGAGAGCGCTGTAACTTGGAGGACAGGATAGGTGAGAACTCAGCCCGCTGTGTC 2340
Db 2281 CTGAGGAGAGCGCTGTAACTTGGAGGACAGGATAGGTGAGAACTCAGCCCGCTGTGTC 2340
Qy 2341 TGTCTAACAAGAGTGCAGATATGCTGCTGTGCTTTCAACACCTACCGGCTTATCCA 2400
Db 2341 TGTCTAACAAGAGTGCAGATATGCTGCTGTGCTTTTCAACACCTACCGGCTTATCCA 2400
Qy 2401 CTGTTTTGATCCATCTCCATCAGAACATCGTGGAGTGCATACCTGCTGACGGTGTAGGT 2460
Db 2401 CTGTTTTGATCCATCTCCATCAGAACATCGTGGAGTGCATACCTGCTGACGGTGTAGGT 2460

Qy 2461 CAGCGTTTGTCTCCTTTGCAATCAAATGGAGTACATCCTGTGTGTTTTCCTTCTCCTGG 2520
Db |||||
Qy 2461 CGGCGGTGTCTCTTTGCAATCAAATGGAGTATGTCCTGTGTCTCTCTCTCTCTCG 2520
Db |||||
Qy 2521 CAGACGCGGGTGTGTGCTGCTTGTGGATGATGCTCTGATAGCCAGGCTGAGGCGG 2580
Db |||||
Qy 2521 CGGACGCGGCTGTGTGCTGCTTGTGGATGATGCTCTGATAGCTCAAGCTGAGGCGG 2580
Db |||||
Qy 2581 CCTTAGAATCTTGTGTGCTCCTCAATGCGGCTCCGTGGCGGAGCGCATGGTATCTCT 2640
Db |||||
Qy 2581 CCCTAGAAACCTGTGTGCTCTCAACGCGGCATCCGTGGCGGGCGCATGGCATCTCT 2640
Db |||||
Qy 2641 CTTTCTTGTGTCTCTCTGCGCGCTGTGTCATTAAGGCGAGGTGCTCTCTGGGCGG 2700
Db |||||
Qy 2641 CTTTCTCTGTGTCTCTCTGCTGTGCTGTATCAAGGCGAGGCTGTCTCTGGGCGG 2700
Db |||||
Qy 2701 CGTATGCTTTTATGCGGTATGCGCGCTGCTCTCTCTACTGCGGTTCACACAGAG 2760
Db |||||
Qy 2701 CATATGCCCTCTACGGCGTATGGCGCTACTCTCTCTCTGCTGGCGTTACACACAG 2760
Db |||||
Qy 2761 CTTACGCTTGGACCGGAGATGGCTGCATGCTGCGGGGGTGGCTTCTGTAGTCTGG 2820
Db |||||
Qy 2761 CATACGCCATGACCGGAGATGGCAGCATGTCGGSAGGCGCGGTTCGTAGGTCTGA 2820
Db |||||
Qy 2821 TATTTCTGACCTTGTCAACCATACTACAAAGTGTTCCTCACTAGGCTCATATGGTGTAC 2880
Db |||||
Qy 2821 TACTCTTGACCTTGTGTCGCACTATAGCTTCTCTCGTAGGCTCATATGGTGTAC 2880
Db |||||
Qy 2881 AATACTTTATCACAGAGCGAGGCGCACATGTCGAAGTGTGGATCCCGCCCTCAACGTT 2940
Db |||||
Qy 2881 AATACTTTATCACAGGCGGAGGCGCACATGTCGAAGTGTGGATCCCGCCCTCAACGTT 2940
Db |||||
Qy 2941 GGGGAGGCGGATGCGCATCTCTCTCACTGCTGCGGTTCATCCAGAGTTAATTTTG 3000
Db |||||
Qy 2941 GGGGGGCGCGATGCGCTCATCTCTCACGTGCGCGATCCACCGAGCTAATCTTTA 3000
Db |||||
Qy 3001 ACATCACCAAACTCTGCTGCGCATACTCGGCCGCTCATGTGCTCCAGGCTGCATAA 3060
Db |||||
Qy 3001 CCATCACCAAACTCTGCTGCGCATACTCGGTCCTCATGTGCTCCAGGCTGTATAA 3060
Db |||||
Qy 3061 CGAGATGCGCTACTTCTGCTGCGCTCAAGGCTCATTCGTGCATGCTAGTTCGAA 3120
Db |||||
Qy 3061 CCAAAGTGCCTACTTCTGCTGCGCGCACAGGCTCATTCGTGCATGCTGTGCGGA 3120
Db |||||
Qy 3121 AAGTGC CGGGGCTCATTTATGTC AAAATGGTCTTATGAGTCTGGGCGGCTGACAGTA 3180
Db |||||
Qy 3121 AGTTGCTGGGGTCATTATGTC AAATGGCTCTCATGAAGTTGGCGCACTGACAGTA 3180
Db |||||
Qy 3181 CGTACGTTTATAACCATCTTACCCACTCGGGACTGGGCCACCGGCGCTACGAGACC 3240
Db |||||
Qy 3181 CGTACGTTTATGACCATCTACCCCACTCGGGACTGGGCCACCGGCGCTACGAGACC 3240
Db |||||
Qy 3241 TTGCGGTGGCGTAGAGCCCGTCTCTCTCGCCATGGAGACCAAGGTCATCACTGGG 3300
Db |||||
Qy 3241 TTGCGGTGGCAGTTGAGCCCGTCTCTCTCTGATATGGAGACCAAGGTTATCACTGG 3300
Db |||||
Qy 3301 GAGCAGACACCGCTGTGGGACATCATCTTGGGTCTACCGTCTCGGCCCGAAAGG 3360
Db |||||
Qy 3301 GGGCAGACACCGCGGTGTGGGACATCATCTTGGGCTTGGCGCTCTCGGCCGAGG 3360
Db |||||
Qy 3361 GGAAGGAGATATTTTGGGACCGGCTGATGTCGAGGGCAAGGTTGGCGACTCTTG 3420
Db |||||
Qy 3361 GGAAGGAGATATCTGGGACCGGCGACAGGCTTTGAAGGGAGGGTGGCGACTCTCG 3420
Db |||||
Qy 3421 CGCCCATCAGGCTACTCTCCAAACAAACGCGGGGCTACTTGGTTGCACTCACTAGCC 3480
Db |||||
Qy 3421 CGCCTATTACGGCTACTCTCCAAACAGACGCGAGGCTACTTGGCTGCATCACTAGCC 3480
Db |||||
Qy 3481 TCACGGCGGGACAAAGAACAGGTTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
Db |||||
Qy 3481 TCACAGGCGGGACAGGAACAGGTTCGAGGGGAGGTCCAAAGTGGTCTCCACCGCAACAC 3540
Db |||||

Qy 3541 AATCTTTCTGGGCACTGTCATCAAGCGCTGTGCTGGACTGTCTACCATGGCGCTGGCT 3600
Db |||||
Qy 3541 AATCTTTCTGGGCACTGTCATCAATGGCGTCAATGGCGTGTGTTGGAATGTCTATCATGGTCCGGCT 3600
Db |||||
Qy 3601 GGAAGACCTTAGCGGTCCAAAAGGTCCAATCACCCAAATGTACACAAATGTAGACCTGG 3660
Db |||||
Qy 3601 CAAAGACCTTAGCGGCCCAAAGGGCCAAATCACCCAAATGTACACAAATGTGGAACAGG 3660
Db |||||
Qy 3661 ACCTCGTGGCTGGCAGCGCCCGCGGGCGCTCCATGACACCATGACAGCTGTGGCA 3720
Db |||||
Qy 3661 ACCTCGTGGCTGGCAGCGCCCGCGGGGGGCTTCTTGACACCATGACCTGCGGCA 3720
Db |||||
Qy 3721 GCTCGGACCTTTACTTTGGTCAAGAGCATGCTGATGTCAATTCGGTGGCGCGGAGGCG 3780
Db |||||
Qy 3721 GCTCGGACCTTTACTTTGGTCAAGAGCATGCGGATGTCAATTCGGTGGCGCGGGGCG 3780
Db |||||
Qy 3781 ACAGCAGGGAAGTCTACTCTCCCCAGGCGCTCTCTACTCTGAAAGGCTCTCTCGGGT 3840
Db |||||
Qy 3781 ACAGCAGGGAAGTCTACTCTCCCCAGGCGCTCTCTACTCTGAAAGGCTCTCTCGGGCG 3840
Db |||||
Qy 3841 GTCCATTTGCTTTGCGCTTCGGGCGACGTGCTGGGCGCTTTCCGGGCTGTGTGCACCC 3900
Db |||||
Qy 3841 GTCCATTTGCTTTGCGCTTCGGGCGACGTGCTGGGCGCTTTTCGGGCTGTGTGCACCC 3900
Db |||||
Qy 3901 GGGGGTTCGGAAGGCGGTGACTTTCATACCCGTTGAGTCTATGGAATCTACCATGGGT 3960
Db |||||
Qy 3901 GAGGGTTCGGAAGGCGGTGACTTTCATACCCGTTGAGTCTATGGAATCTACCATGGGT 3960
Db |||||
Qy 3961 CTCGGTTCACAGACAACTCAACCCCGCGGCTGTACCGCAGACATTCGAAGTGGAC 4020
Db |||||
Qy 3961 CCGCGTTCACAGACAACTCGTCCCTCGGCGCTACCGCAGACATTCGAAGTGGCG 4020
Db |||||
Qy 4021 ATCTGACGCTCTCTACTGCGAGGCGGCAAGAGCACAAAGTCCGGCTGCGTATGCAGCC 4080
Db |||||
Qy 4021 ATCTGACGCTCTCTACTGCGAGGCGGCAAGAGCATTAAGTTCGGGCTGCGTATGCAGCC 4080
Db |||||
Qy 4081 AAGGTTCAAGGTCTGCTGCTGAACCCGCTCGTTGCGCGCACCTTAGGGTTTGGGGCGT 4140
Db |||||
Qy 4081 AAGGTTAAGGTCTGCTGCTGAACCCGCTCGTTGCGCGCACCTTAGGGTTTGGGGCGT 4140
Db |||||
Qy 4141 ATATGTC AAGGCA CAGGTCATGACCCCTTAACATCAGAACTGGGGTGAAGCAATACCA 4200
Db |||||
Qy 4141 ATATGTC AAGGCA CAGGTCATGACCCCTTAACATCAGAACTGGGGTGAAGCAATACCA 4200
Db |||||
Qy 4201 CGGGGCTCCATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4260
Db |||||
Qy 4201 CGGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4260
Db |||||
Qy 4261 GGGGCGCTGATGACATCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 4320
Db |||||
Qy 4261 GGGGCGCTGATGACATCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 4320
Db |||||
Qy 4321 TGGGATCGGCAAGTCTGGAACAGGAGAGCGGTGAGGAGGCGGCTGCTGCTGCTGCTG 4380
Db |||||
Qy 4321 TGGGATCGGCAAGTCTGGAACAGGAGAGCGGTGAGGAGGCGGCTGCTGCTGCTGCTG 4380
Db |||||
Qy 4381 CCACGCTTACACCTTCGGGATCGGTTACCGTCCACACCCCAATATCGAGGAATAGGCC 4440
Db |||||
Qy 4381 CCACGCTTACACCTTCGGGATCGGTTACCGTCCACACCCCAATATCGAGGAATAGGCC 4440
Db |||||
Qy 4441 TGTCCAAATGAGAGATCCCTTTCTATGCAAGGCTATCCCATTTGAGGCGCATCAAGG 4500
Db |||||
Qy 4441 TGTCCAGCATGAGGAATCCCTTTTATGCAAGGCTATCCCATCGAGACCATCAAGG 4500
Db |||||
Qy 4501 GGGGAGGCTCTCATTTTCTGCAATTCGAAGGAATGTCGAGCTCGCGGCAAGC 4560
Db |||||
Qy 4501 GGGGAGGCTCTCATTTTCTGCAATTCGAAGGAATGTCGAGCTCGCGGCAAGC 4560
Db |||||
Qy 4561 TGACAGGCTCGGACTGACCTGTAGCATATTCGCGGGCTTGTATGTCGCTGCTATAC 4620
Db |||||
Qy 4561 TGTCGGGCTCGGACTCAATGCTGAGCATATTCGCGGGCTTGTATGTCGCTGCTATAC 4620
Db |||||
Qy 4621 CGCCTATCGGAGAGCTGCTGCTGTCGCAACAGACGCTCTTAATGACGGGTTTCAACGGCG 4680
Db |||||

Qy	6841	ACGTAA	CAGTGC	TTACTTC	CATGCT	CACGAT	CCCTCC	CAATTA	CAGCAG	AAGACG	GGCTA	6900
Db	6841	ACGTAG	CAGTGC	TCTCACT	TCCATG	CTCACC	GCACCC	TCCCACT	ATACGG	CGGACG	GGCTA	6900
Qy	6901	AGCGTAG	GTGGCT	TAGAGG	GTCTCC	CCCTCT	TTTAGC	CAGCTCAT	CAGCTAG	CAGCTTGT	6960	
Db	6901	AGCGTAG	GTGGC	CAGGGA	TCTCCC	CCCTCT	TTTGGC	CAGCTCAT	CAGCTAG	CAGCTGT	6960	
Qy	6961	CTGGC	CTCTTT	GAAGCG	CACATG	CACTACC	CAACAT	GACTCCC	CGGACG	CTGACTCA	7020	
Db	6961	CTGGC	CTTCTTT	GAAGGCA	CAATGCA	TCTAC	CGTCA	TGACTCCC	CGGACG	CTGACTCA	7020	
Qy	7021	TCGAGG	CCAACT	CTTGTG	CGGCGAG	GATGGG	CGGAAACAT	CACTCG	CGTGGAGT	CGAGTCA	7080	
Db	7021	TCGAGG	CCAACT	CTCTGT	GGCGG	CAGAGATGG	CGGGAACAT	CACCCG	GTGGAGT	CGAGTCA	7080	
Qy	7081	AGATA	AGGTAG	TAAAT	TCTG	GAATCT	TTTGAAC	CGGCTT	CAACGG	GAGGGATG	AGAGG	7140
Db	7081	AAAATA	AGGTAG	TAAAT	TTTGA	CTCTTC	GAGCGCT	CCAAG	CGGAGG	AGGATG	AGAGG	7140
Qy	7141	AGATAT	CCGTG	CGGCGG	AGATCT	CTGCG	AAATCC	CAGGA	AGTTCC	CTCAG	CGTTGCC	7200
Db	7141	AAGTAT	CCGT	TCGCGG	CGAGAT	CTCGG	AGGTCC	CAGGA	AATTCCT	TCGAG	CGATGCC	7200
Qy	7201	TATGGG	CACG	CCGGA	CTACA	ATCTCT	CCACTGT	GTAG	AGTCTCG	GAAGGAC	CCCGACTACG	7260
Db	7201	TATGGG	CACG	CCGGA	TTC	AACTCT	CACTGT	TAG	AGTCTCT	GGAAGG	ACCCGACTACG	7260
Qy	7261	TCCCT	CCGGT	GGTAC	CGGAT	TCGCA	CTAC	CAAGG	GTCTCT	CCAATAC	CACTC	7320
Db	7261	TCCCT	CCAGT	GGTAC	CGG	TGTC	CAATG	TCCT	CAAGG	CCCTC	CGATAC	7320
Qy	7321	CAGG	GAAG	AGG	ACG	GTGTCT	GTAC	AGAA	TCCAT	GTGCTT	CTG	7380
Db	7321	CACGG	AGAA	AGG	ACG	GTGTCT	GTG	CA	AACT	CTAC	CGTGTCT	7380
Qy	7381	TCGG	CACT	TA	AG	ACCT	TCGG	TAG	CTCG	CGCTT	GTATAG	7440
Db	7381	TCGG	CCAA	AG	ACCT	TCGG	AGCTT	CG	AA	TGCT	CGCCGT	7440
Qy	7441	CCCT	TCCT	GAC	TCG	CCCT	CCGAC	AGG	TGCA	AAAGG	ATCCG	7500
Db	7441	CTCT	CTG	AC	CA	CCCT	CGAC	GC	GGAT	CCG	AGTTG	7500
Qy	7501	CCAT	GCCCC	CCCTT	GA	AGG	AGCCG	GGG	ACCC	CGATCT	CAGG	7560
Db	7501	CCAT	GCCCC	CCCTT	GA	GGG	AGCCG	GGG	ATCC	CGG	ATCT	7560
Qy	7561	CCGT	GAGT	GAGG	CGCT	TAG	GAGT	CTCG	TGCT	CAAT	GTCTAT	7620
Db	7561	CCGT	TA	AG	CG	AGG	AGGCT	TAG	TAG	AC	CTG	7620
Qy	7621	CGCG	CCCT	GTAT	CAC	CCCAT	TGCG	CTCG	GAGAA	ATAAG	CTGCC	7680
Db	7621	CGCG	CCCT	GTAT	CAC	CCCAT	TGCG	CTCG	GAGG	AAACCA	AGCTG	7680
Qy	7681	ACT	CTTT	GCTG	CGTCA	CA	CAAC	ATGTT	CTAC	GC	CAAG	7740
Db	7681	ACT	CTTT	GCT	CGT	CA	CAAC	AT	TTG	CTAT	G	7740
Qy	7741	GGC	AGA	AGG	TAC	CTTT	TG	AC	AG	ATTG	CA	7800
Db	7741	GGC	AG	AG	AGG	TAC	CTTT	TG	AC	AG	ATTG	7800
Qy	7801	TCA	AGG	AG	ATGA	AGG	CGCT	CC	AC	AGT	TTA	7860
Db	7801	TCA	AGG	AG	ATGA	AGG	CGG	TCC	AC	AGT	TTA	7860
Qy	7861	CCT	GCA	AG	CTGA	CG	CCCC	CA	ATTT	CGG	CTA	7920
Db	7861	CCT	GTA	AG	CTGA	CG	CCCC	CA	ATTT	CGG	CTA	7920

Qy	7921	TCCGGAACCTATCCAGCAGGGCCGTTAA	CCACATCCGCTCGTGTGGAGGACTTGTCTGG	7981	
Db	7921	TCCGGAACCTATCCAGCAGGGCCGTTAA	CCACATCCGCTCGTGTGGAGGACTTGTCTGG	7981	
Qy	7981	AAGACACTGAAAACACCAATTGACACCA	CACCATCATGCAAAAAGTGAGTTTCTGGGTCC	8040	
Db	7981	AAGACACTGAGACACCAATTGACACCA	CACCATCATGCAAAAAGTGAGTTTCTGGGTCC	8040	
Qy	8041	AACCAGAGAGGGAGGCGCAAGCCAGCT	TATCGTATTCCAGACCTCGGAGTTTC	8100	
Db	8041	AACCAGAGAGGGGCGCGCAAGCCAGCT	TATCGTATTCCAGATTTGGGGTTTC	8100	
Qy	8101	GTGTATCGCAGAAAGATGGCCCTTTA	CGAGTGTCTCCACCTTCCTCAGCCGCGTATGG	8160	
Db	8101	GTGTGTGCGAGAAAATGGCCCTTTA	CGATGTGTCTCCACCTTCCTCAGCCGCGTATGG	8160	
Qy	8161	GCTCCTCATACGGATTTCANA	TATCCCCCCAAGCAGCGGTGCGAGTTTCTGGTCAATACCT	8220	
Db	8161	GCTCTTCATACGGAATCCAA	TATCTCTCTGGACAGCGGTGCGAGTTTCTGGTCAATACCT	8220	
Qy	8221	GGAAATCAAAAGAAATGCCCTAT	TGGGCTTCATATGACACCCGCTGTTTGA	8280	
Db	8221	GGAAAGCGAAGAAATGCCCTAT	TGGGCTTCGCATATGACACCCGCTGTTTGA	8280	
Qy	8281	TCAC	TGAGAGTGACATTCGTGTGGAGAGTCA	ATTATTCACCAATGTGTGACTTTGGCCCCCG	8340
Db	8281	TCAC	TGAGATGACATCCGTGTTGAGAGTCA	ATCTACCAATGTGTGACTTTGGCCCCCG	8340
Qy	8341	AGCCGACAGAGGCATNAAGT	TCGCTCACAGAGCGGCTTTACATCGGGGTTC	8400	
Db	8341	AAGCCGACAGAGGCATNAAGT	TCGCTCACAGAGCGGCTTTACATCGGGGGCCCCCTGACTA	8400	
Qy	8401	ACTCAAAAGGGCAGAACTCGGGT	TATCGCGGTGCGCGCAAGTGGCGTGCTGACACGACTA	8460	
Db	8401	ATTCTAAAGGGCAGAACTCGCGT	ATCGCGGTGCGCGCGAGCGGTGTACTGACGACCA	8460	
Qy	8461	GCTGCGGTAA	TACCCCTCACATGTTACTTTGAAGCCAC	TGACGAGCTGTGCGAGCTGCAAAAGC	8520
Db	8461	GCTGCGGTAA	TACCCCTCACATGTTACTTTGAAGCCCGCTGCGGCTGTGCGAGCTGCGAAGC	8520	
Qy	8521	TCGAGGACTCGACGATGCTCGT	GAAACGGAGACGACTTGTGCTGTATCTGTGAAAGCGCGG	8580	
Db	8521	TCCAGGACTCGACGATGCTCGT	ATGCGGAGACGACCTTGTGCTGTATCTGTGAAAGCGCGG	8580	
Qy	8581	GNACCCAGGAGGATGCGCGGCC	CTACGAGCTTCAGGAGGCTATGACTAGGTATTCCG	8640	
Db	8581	GGACCCAAAGAGGACGAGCGAGC	CTACGAGGCTTCAGGAGGCTATGACTAGATATCTCTG	8640	
Qy	8641	CCCCCCCCGGGGATTC	CGCCCCCAACCAAGATACGACCTGGAGCTGATAACATCATGTTCTCT	8700	
Db	8641	CCCCCCCCGGGGACCGCCCCA	ACCAAGATACGACTTGGAGTTGATAACATCATGCTCCT	8700	
Qy	8701	CCAAATGTGT	CAGTGGCGACGATGCTGTGGCAAAAGGGTAT	ACTACCTCAACCGGTGACC	8760
Db	8701	CCAAATGTGT	CAGTGGCGACGATGCTGTGGCAAAAGGGTGTACTATCTCAACCGGTGACC	8760	
Qy	8761	CCACACCCCGCTTG	CACGGGCTCGGTGGGAGACAGCTAGACACACTCCAATCAACTCTT	8820	
Db	8761	CCACACCCCGCTTG	CGCGGGGCTCGGTGGGAGACAGCTAGACACACTCCAAGTCAAAATCTCT	8820	
Qy	8821	GGCTAGGCAATAT	CATCATGTATGCGGCCACCCCTATGGGCAAGGATGATTCTGATGACTC	8880	
Db	8821	GGCTAGGCAATAT	CATCATGTATGCGGCCACCTTGTGGGCAAGGATGATCCTGATGACTC	8880	
Qy	8881	ACTTTTTCT	CCATCCTTCTAGCTCAAGAGCAACTTGAAAAAGCCCTGGATTTGT	8940	
Db	8881	ATTTCTTCT	CCATCCTTCTAGCTCAGGAACAACTTGAAAAAGCCCTAGATTGT	8940	
Qy	8941	ACGGGGCTT	TGCTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTGACGCTCCATG	9000	
Db	8941	ACGGGGCTT	TGCTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTCAACGACTCCATG	9000	
Qy	9001	GTCTTAGCGAATTT	TACACTCCACAGTTTACTCTCCAGGTGAGATCAATATAGGGTGGCTTTCAT	9060	


```
Db 9001 GCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGTGAGATCAATAGGGTGGCTTCAT 9060
Qy 9061 GCCTTAGGAACTTTGGGTACCAACCTTGGAACTTGGAGACATCGGGCCAGAAAGTGTC 9120
Db 9061 GCCTTAGGAACTTTGGGTACCGCCCTTGGAGTCTGGAGACATCGGCCAGAAAGTGTC 9120
Qy 9121 GCCTTAAGCTACTGTCCAGGGGGGAGGGCCGCCACTTGTGGCAGATACCTCTTTAACT 9180
Db 9121 GCCTTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTCAACT 9180
Qy 9181 GGGCAGTAGGACCAAGCTTAAACTCACTCCATCCCGGCCGCTCCAGCTCGACTGTG 9240
Db 9181 GGGCAGTAGGACCAAGCTCAAACTCACTCCATCCCGGCTGGCTCCAGTTGGATTTAT 9240
Qy 9241 CTGGCTGGTTTCGTGGTGTACAGCGGGGGAGACATATACAGCCTGTCTCGTGCCC 9300
Db 9241 CAGCTGGTTTCGTGGTGTACAGCGGGGGAGACATATACAGCCTGTCTCGTGCCC 9300
Qy 9301 GACCCGCTGGTTTCGTGGTGTCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9360
Db 9301 GACCCGCTGGTTTCATGTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9360
Qy 9361 TCCCAACCGATGAACGGGGAGCTAACCTCCAGGCCCTTAAGCC-----ATTTC 9410
Db 9361 TCCCAACCGATGAACGGGGAGCTAAACACTCCAGGCCCAATAGGCCATCTCTGTTTTTC 9420
Qy 9411 CTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9470
Db 9421 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9480
Qy 9471 TTTTCTTTCTTTTTCCTTTCTTTAAATGGTGGCTCCATCTAGCCCTTAGTCACGGCTAGC 9530
Db 9481 TCCTCTTTTTTTCCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
Qy 9531 TGTGAAAGTCCGTGAGCGGATGACTGCAGAGAGTCTGATCTAGCCCTTAGTCAGAGAT 9590
Db 9541 TGTGAAAGTCCGTGAGCGGCTTGACTGCAGAGAGTGTGATCTAGCCCTCTCTGCGAGT 9600
Qy 9591 CATGT 9595
Db 9601 CAAGT 9605
```

```
RESULT 4
US-10-475-024-17
; Sequence 17, Application US/10475024
; Publication No. US20040219545A1
; GENERAL INFORMATION:
; APPLICANT: Rando, Robert F.
; APPLICANT: Welch, Ellen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING SMALL MOLECULES THAT BIND SPECIFIC RNA
; TITLE OF INVENTION: STRUCTURAL MOTIFS
; FILE REFERENCE: 10589-007-999
; CURRENT APPLICATION NUMBER: US/10/475,024
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/282,965
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 9353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-475-024-17
```

```
Query Match 85.3%; Score 8185; DB 8; Length 9353;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 8666; Conservative 0; Mismatches 685; Indels 60; Gaps 1;
Qy 14 TGGGGGGGACACTCCACCATGAATCACTCCCTGTGAGGAACTACTGTTTCACGAGAA 73
Db 2 TGGGGGGGACACTCCACCATGATCACTCCCTGTGAGGAACTACTGTTTCACGAGAA 61
```

```
Qy 74 AGCGTCTAGCATGGGTAGTTAGTGTCTGTCAGAGCTCCAGGACCCCTCCCGGG 133
Db 62 AGCGTCTAGCATGGGTAGTTAGTGTCTGTCAGAGCTCCAGGACCCCTCCCGGG 121
Qy 134 AGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAAATGCCAGGACGACCGGGTCC 193
Db 122 AGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAAATGCCAGGACGACCGGGTCC 181
Qy 194 TTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGGGTGCCCCCGCGAGACTGCTAG 253
Db 182 TTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGGGTGCCCCCGCGAGACTGCTAG 241
Qy 254 CCGAGTAGTGTGGGTGCGGAAAGGCTTGTGTACTGTCTGATAGGTTGCTTGCAGTG 313
Db 242 CCGAGTAGTGTGGGTGCGGAAAGGCTTGTGTACTGTCTGATAGGTTGCTTGCAGTG 301
Qy 314 CCGCGGAGGTCTCTGTAGACCGGTGACCAATGAGCAGCAATCTCTAAACCTCAAGAAAAAC 373
Db 302 CCGCGGAGGTCTCTGTAGACCGGTGATCATGAGCACAATCTCTAAACCTCAAGAAAAAC 361
Qy 374 CAAACGTAAACCAACCCGCGCCACAGAGCGTCAAGTTCCCGGGGGTGGTGCAGATCGT 433
Db 362 CAAACGTAAACCAACCCGCGCCACAGAGCGTCAAGTTCCCGGGGGTGGTGCAGATCGT 421
Qy 434 TGGTGGAGTTTACTGTTGCGCGGAGGGGCCCCAGGTTGGGTGGTGGCGCGACTAGGAA 493
Db 422 TGGTGGAGTTTACTGTTGCGCGGAGGGGCCCCAGGTTGGGTGGTGGCGCGACTAGGAA 481
Qy 494 GGCCTCCGAGCGTCCCAACTCGTGGAAAGGCGCAACCTATATCCAAAGGCTCCGCGACC 553
Db 482 GACTTCGAGCGGTGCAACCTCTGAGAGCGCAACCTATATCCCAAGGCTCCGCGGCC 541
Qy 554 CGAGGGCAGGGCTGGGCTCAGCCCGGGTACCCTTGGCCCTCTATGGCAATGAGGCGCT 613
Db 542 CGAGGGTAGGACTGGGCTCAGCCCGGGTACCCTTGGCCCTCTATGGCAACGAGGGTAT 601
Qy 614 GGGGTGGGAGGATGGTCTGTACCCCGGGGCTCCCGGCTAGTTGGGGCCCCACGGA 673
Db 602 GGGGTGGGAGGATGGTCTGTACCCCGGGGCTCTCGGCGCTAGTTGGGGCCCCACAGA 661
Qy 674 CCCCAGGCTAGGTCCGTAACCTTGGTAAAGTTCATCGATACCTTACATCGGCTTCGC 733
Db 662 CCCCAGGCTAGGTCCGTAACCTTGGTAAAGTTCATCGATACCTTACATCGGCTTCGC 721
Qy 734 CGATCTCATGGGGTACATTCGCTCGGCGGCCCTTAGGGGGCGCTGCCAGGGGCTT 793
Db 722 CGACCTCATGGGGTACATTCGCTCGGCGGCCCTTAGGGGGCGCTGCCAGGGGCTT 781
Qy 794 GGCACAGGTGTCCGGGTTCTGAGGACGGCGTGAATATGCAACAGGGAACCTTGCCTCG 853
Db 782 GGCACATGGTGTCCGGGTTCTGAGGACGGCGTGAATATGCAACAGGGAATCTTGCCTCG 841
Qy 854 TTGCTCTTTCTCTATCTTCTCTTGGCTCTGCTCTGCTGTTGACCACTCCAGCTTCGC 913
Db 842 TTGCTCTTTCTCTATCTTCTCTTGGCTCTGCTCTGCTGTTGACCACTCCAGCTTCGC 901
Qy 914 TTATGAAGTGCAGAACGCTGTCGGGATATACCATGTCAAGAACGACTGCTCCAACTCAAG 973
Db 902 TTACGAGTGGCGCAACGCTGTCGGGATATACCATGTCAAGAACGACTGCTCCAACTCAAG 961
Qy 974 CATTTGTATGAGGACGGGACGTGATCATGCATATCTCCGGGTGGTGGCTCTGTTCAT 1033
Db 962 TATTGTATGAGGACGGGACGTGATCATGCACACCCCGGGTGGTGGCTCTGCTCG 1021
Qy 1034 GGAGGGTAAACGCTCCCGTTCGTGGGTAGCGTCACTCCACAGCTCCGGGCGAGAAATGC 1093
Db 1022 GGAGAGTAAATTTCTCCCGTTCGTGGGTAGCGTCACTCCACAGCTCCGGGCGAGAAATGC 1081
Qy 1094 CAGCGTCCCACTACGACAATACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 1153
Db 1082 CAGCATCCCCACCGACAATACGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 1141
```

QY 1154 CTGCTCCGCTATGTACGTGGGGATCTCTCGGGATCTATTTCCTCGTCTCCAGCTGT 1213
DB 1142 CTGTTCCGCTATGTACGTTGGGGATCTCTGGGGATCCGTTTCTCGTCTCCAGCTGT 1201
QY 1214 CACCTTCTCGCCTCCGCGCATGAGACAGTCAGAGACTGCAACTGCTCAATCTATATCCCGG 1273
DB 1202 CACCTTCTCACCTCCGCGTATGAGACGGTACAAGATTGCAATTGCTCAATCTATATCCCGG 1261
QY 1274 CCATGTATCAGGTACCGCATGGCTTGGGATATGATGATGAACCTGCTCACTACAAACAGC 1333
DB 1262 CCACGTATCAGGTACCGCATGGCTTGGGATATGATGATGAACCTGCTCACTACAAACAGC 1321
QY 1334 CCTAGTGTGTGCGAGTTGCTCCGGATCCCAAGCTGTCTGTGGACATGCTGGCGGGGC 1393
DB 1322 CCTAGTGTATCGCAGCTACTCCGGATCCCAAGCGCTGTGTGGACATGCTGGCGGGGC 1381
QY 1394 CCACCTGGGAGTCTGGCGGGCTTGGCTTATTTCCATGTGTAGGAACTGGGCTTAAGGT 1453
DB 1382 CCACCTGGGAGTCTTAGCGGGCTTGGCTTATTTCCATGTGTGGGAACTGGGCTTAAGGT 1441
QY 1454 TCTGATTGGCGCTACTCTTTGGCGGGTGAACGGGGAGACCCACAGCGGGAGGGT 1513
DB 1442 CTTGATTGTGATGCTACTCTTTGCTGGCGTTGACGGGCACACCCACGTGACAGGGGGAAG 1501
QY 1514 GGCGGGCCACACCACTCCGGGTTCAAGTCCCTTTTCTCATCTGGGGGTCTTCAGAAAAT 1573
DB 1502 GGTAGCTCCAGCACCCAGAGCTCGTCTCGGCTCTCAAGAGCCCATCTCAGAAAAT 1561
QY 1574 CCAGCTTGTGAATACCAACGGCAGTGGGCACATCAACAGGACTGCCCCATAATGCAATGA 1633
DB 1562 CCAACTCTGTGAACCAACAGGAGCTGGGCACATCAACAGGACCGCTCTGAATTGCAATGA 1621
QY 1634 CTCCTTCCAACTGGGTTCTTTGGCGCGTGTGTTACGCACACAGTTCACTGCTCCGG 1693
DB 1622 CTCCTTCCAACTGGGTTCAATTGCTGCGTGTGTTACGCACACAGTTCAACGCGCTCCGG 1681
QY 1694 GTGCCCGAGCCATGCGGAGCTGCGGCCCATTCGACTGTTGCGCCAGGGGTGGGGCCC 1753
DB 1682 GTGCCAGAGCCGATGGCTAGCTGCGGCCCATCGATGAGTTTCGTCAGGGGTGGGGTCC 1741
QY 1754 CATCACTTATATAAGCCTTAAACAGTTCGGATCAGAGGCTTATTCCTGGCATTAACGCGCC 1813
DB 1742 CATCACTCATGATATGCTCTGAGAGCTCGGACACAGAGGCCATATGCTGGCATTAACGCGCC 1801
QY 1814 TCGACGCTGTGTGCTGTAACCGCTGCGAGGTGTGTCAGTGTATTGTTTCAACCC 1873
DB 1802 TCGACGCTGCGGGATCGTGCCTGCGCTGCGAGGTGTGTCAGTGTATTGCTTCACTCC 1861
QY 1874 AAGCCCTGTTGTGGGGACCAACCGATCGTTTCCGCTGTCCTACGTATAGCTGGGGGA 1933
DB 1862 GAGCCCTGTTGTAGTGGGACACCGATCGTTTCCGCGCTCTACGTATAGCTGGGGGA 1921
QY 1934 GAATGAGACAGAGTGATGCTCCTCAACACAGCGTCCGCCAACAGGCAACTGTTTCGG 1993
DB 1922 GAATGAGACAGAGTGCTGCTACTTATAGCAACACGCGCGCGCTCAAGGCAACTGTTTGG 1981
QY 1994 CTGTACATGATGAATAGTACTGGGTTCACTAAGACGTGCGAGGTCCCGCTGTAAAT 2053
DB 1982 GTGCACTGGATGAACAGACACTGGGTTTCAACAGACGTGCGGGGGCCCTCGCTGCAACAT 2041
QY 2054 CGGGGGGTGCGGTAAACCGCACTTGATCTGCCCCAGGACTGTTCCGGAAGCAACCCGA 2113
DB 2042 CGGGGGGTGCGGCAACCAACCTTGTGTGTCGCCACGAGTTGTTCCGGAAGCAACCCGA 2101
QY 2114 GGCTACTTACAAAATGTGGTTCGGGGCCCTGGTTGACACCTAGGTGCTTAGTACTGA 2173
DB 2102 GGCCACTTACAAAAGTGTGGTTCGGGGCCCTGGTTGACACCCAGGTGATGTTGACTA 2161
QY 2174 CCCATACAGGCTTGGCACTACCCCTGCACTCTCAATTTTTCATCTTTAAGGTTAGGT 2233
DB 2162 CCCATACAGGCTTGGCACTACCCCTGCACTCTGTTAACTTTTACCGTCTTTAAGGTCAGGT 2221
QY 2234 GTATGTGGGGGCGTGGAGCACAGGCTCAATGCGGCATGCAATTGGACTCGAGGAGAGCG 2293

DB 2222 GTATGTGGGGGCGTGGAGCACAGGCTCAATGCTGCATGCAATTGGACTCGAGGAGCG 2281
QY 2294 CTGTAACTTTGGAGACAGGNTAGTTCAGACTCAGCCGCTGCTGTCTTACAAACAGA 2353
DB 2282 CTGTGACTTTGGAGACAGGNTAGTTCAGAACTCAGCCGCTGCTGTCTTACAAACAGA 2341
QY 2354 GTGCAGACTACTGCGCTGTGCTTTTCCACCCCTACCGGCTTATCCACTGCTGTTGATCCA 2413
DB 2342 GTGCAGACTACTGCGCTGTGCTTTTCCACCACTACCGGCTGCTCCACTGCTGTTGATCCA 2401
QY 2414 TCTCCATCAGAACTCTGTGACGTGCAATACCTGTACGGTGTAGGGTCAGGTTTGTCTC 2473
DB 2402 TCTTCAACGGAACTCTGTGACGTGCAATACCTGTACGGTGTAGGGTCAGGTTTGTCTC 2461
QY 2474 CTTTGGCAATCAATGGGAGTACATCTGTTGCTTTTCTTCTCTGCGACAGCGCGGT 2533
DB 2462 CTTTGGCAATCAATGGGAGTATATCTGTTGCTTTTCTTCTTCTGCGGAGCGCGGT 2521
QY 2534 GTGTGCTGCTGTGTGGATGATGCTGTGATAGCCAGGCTGAGCGCGCTTAGAGAACTT 2593
DB 2522 CTGTGCTGCTGTGTGGATGATGCTGTGATAGCCAGGCTGAGGCGCACCTTAGAGAACTT 2581
QY 2594 GGTGCTCTCAATCGCGCTCGTGGCGGAGCGCATGTTTCTTCTCTTCTTCTGTTGT 2653
DB 2582 GGTGCTCTCAATCGCGCTCTGTGGCGGAGCGCATGGCCTTCTCTCTTCTCTGTTGT 2641
QY 2654 CTTCTGCGCGCTGTGTACATTAAAGGCGAGGCTGCTCTGCGGGGCGGTATGCTTTT 2713
DB 2642 CTTCTGCGCGCTGTGTACATTAAAGGCGAGGCTGCTCTGCGGGGCGGTATGCTTCT 2701
QY 2714 TGGCGTATGGCGCTGCTCTGCTCTTACTGCGGCTTACCACACGAGCTTACGCGCTTGA 2773
DB 2702 TGGCGTATGGCGGTTGCTCTGCTCTTCTGCTGCGCTTACCACACGAGCTTATGCGATGA 2761
QY 2774 CCGGAGATGCTCATCGTGGGGGTGCGGTTCTTGTAGGTCTGTGPAITCTTGACCTT 2833
DB 2762 CCGAGAGATGCTCATCGTGGGAGGCGGTTTGTAGGTCTGTGTAATCTTGACCTT 2821
QY 2834 GTACCATATCAAAAGTGTTCATAGGCTCATATGTTGGTGTACAAATCTTATCAC 2893
DB 2822 GTACCATATCAATAGGTTTCTCGTAGGCTCATATGTTGGTGTACAAATCTTATCAC 2881
QY 2894 CAGAGCGAGCGCACATGCAAGTGTGGTCCCGCTCAACGTTCCGGGAGGCGCGGA 2953
DB 2882 CAGAGCGAGCGCACATGCAAGTGTGGTCCCGCTCTCAATGTTCCGGGAGGCGCGGA 2941
QY 2954 TGCCATCATCTCTCATCGTGGGTTTCTCCAGAGTTAAATTTTGAATCATCAACAACT 3013
DB 2942 TGCCATCATCTCTCATCGTGGGTTTCTCCAGAGCTAATCTTTGACATCAACAACT 3001
QY 3014 CTGCTGCGCATACTCGGCGCGCTCATGTTGCTCCAGCTGGCATAGAGAGTCCGTA 3073
DB 3002 CTGCTGCGCATACTCGGCTCGGCTCATGTTGCTCCAGCTGGCATAGAGTCCGTA 3061
QY 3074 CTTGCTGCGCTCAAGGGCTCATTTGTCATGATGTTAGTGGCAAAAGTCCGCGGG 3133
DB 3062 CTTTGTACCGCTCAGGGGCTCATCCGTCATGATGTTAGTGGCAAGTTCGCTGGAG 3121
QY 3134 TCATTATGTCCAAATGTTCTTTCATGAAGCTGGGCGCTGACAGGTACGTAGTTTATAA 3193
DB 3122 CCATATGTCCAAATGGCTTTCATGAAGTGGCGCGCTGACAGGTACGTAGTATGA 3181
QY 3194 CCATCTTACCCACTGCGGAGCTGGGCGGCTGAGGCGCTTACGAGCTTGGGTCGCGGT 3253
DB 3182 CCATCTTACCTGCGGAGTGGGCGGCTGAGGCGCTTACGAGCTTGGGTCGCGGT 3241
QY 3254 AGAGCGGCTGCTTCTCCGCGATGGAGACCAAGTCTACCTGCGGAGAGACACCGC 3313
DB 3242 AGAGCGGCTGCTTCTCTGACATGGAGCTTAACTCATCATCTGCGGGGAGACACCGC 3301
QY 3314 TGCGTGTGGGAGACTCATCTTGGGTCTACCGCTCTCCCGCCGAAGGGGAAGGATATT 3373

Db	3302	GGCGTGTGGGGA	CATCATCTCGGGGTCTACAGTCTTCGGCCCGAAGGGGAGAGGATACT	3361
Qy	3374	TTTGGGACCGGCTGTAGTCTCGAAGGGCAAGGGTGGCGACTCTTTTCGGCCCATCACGGC	3433	
Db	3362	TCTAGGACCGGCGATAGTTTGGAGAGCAGGGGTGGCGCTCTTTGGCCTATCACGGC	3421	
Qy	3434	CTACTCCCAACAAACGGGGGGCGTATCTTGTTGTCATCATCTAGCTCATCAGGCCGGGA	3493	
Db	3422	CTATTCCCAACAAACGGGGGCGCTGTCTGCTGTATCATCTATGACTCATCAGGTCCGGA	3481	
Qy	3494	CAGAACCAGGTCGAAGGGGAGGTTCAAGTGGTTTCTACCGCAACAACATCTTTCTCTGGC	3553	
Db	3482	CAAGAACCAAGTCGATGGGAGGTTCAAGTGGTCTCTCACCGCAACCAATCTTTCTCTGGC	3541	
Qy	3554	GACCTGCATCAACGGCGGTGTCTGGACTGTCTACCATGGCGCTGGCTCGAAGACCCCTAGC	3613	
Db	3542	GACCTGGCTCATGGCGGTGTGGACCGTCTACCATGGTGGCGCTCGNAGACCCCTGGC	3601	
Qy	3614	CGGTCCAAAAGTCCAAATCAACCAAAATGTATACCAATGTATAGACTGGACCTCGTCGGCTG	3673	
Db	3602	CGGCCGAAGGTCATCAACCAAAATGTATACCAATGTATAGACAGGACCTCGTCGGCTG	3661	
Qy	3674	GCAGGCGCCCCCGGGCGCGCTCCATGACACCAATGCAGCTGTGGCAGCTCGACCTTTA	3733	
Db	3662	GCGGCGCCCCCGGGCGCGCTCCATGACACCGTGCACTGGCGGAGCTCGACCTTTA	3721	
Qy	3734	CTTGGTCAAGAGACATGCTGATGTATCTCCGTTGGCGCGCGGAGGCGACAGCAGGGGAAG	3793	
Db	3722	CTTGGTCAAGGCATGCTGATGTCTGTTCCGTTCCGCGCGGGGCGACAGCAGGGGGAG	3781	
Qy	3794	TCTACTCTCCCCAGGCGCGCTCTCTACTCTGAAGGCTCTCCGGTGGTCCATTTGCTTTG	3853	
Db	3782	CCTGCTTTCCCCAGGCGCATCTCTACTCTGAAGGCTCTCCGGTGGACCATGCTGTTG	3841	
Qy	3854	CCCTTCGGGCGACGTCGTGGGCGCTTTCGGGGCTGTGTGCAACCGGGGGTGCAGAA	3913	
Db	3842	CCCTTCGGGCGACGTTGTAGGCATCTTTCGGGGCTGTGTGTGCAACCGGGGGTTCGAA	3901	
Qy	3914	GGCGGTGGACTTACATACCCGTTGAGTCTATGGAACCTACCATCGGCTCTCCGGTCTTCAC	3973	
Db	3902	GGCGGTGGACTTATATACCCGTTGAGTCTATGGAACCTACCATCGGCTCTCCGGTCTTCAC	3961	
Qy	3974	AGACAACTCAACCCCGCGCTGTACCGCAGACATTCAAAGTGGCACATCTGCACGCTCC	4033	
Db	3962	AGACAACTCATCCCTCTCGGCGGTACCGCAACATTCAAAGTGGCACATTTACAGCTCC	4021	
Qy	4034	TACTGGCAGCGGCAAGAGCACAAAGTGCCGGCTGGTATGCGCCCAAGGGTACAAGGT	4093	
Db	4022	CACCTGGCAGCGGCAAGAGCACAAAGTGCCGGCTGCATATGCAAGCCCAAGGTACAAGGT	4081	
Qy	4094	GCTCGTCTGAACCGGTCCGTTGGCGCACCTTAGGGTTTGGGCGGTATATGTCACAGGC	4153	
Db	4082	GCTCGTCTGAACCGGTCCGTTGGCGCACATGCGGCTTTGGAGCGGTATATGTCACAGGC	4141	
Qy	4154	ACAGGTATCGACCTTAACATCAGAACTGGGGTAAGGACCATTAACACGGCGGCTCCAT	4213	
Db	4142	ACATGGCATCGAGCGCTAATCAGAACTGGGGTAAGGACCATTAACACGGCGGCGCCAT	4201	
Qy	4214	TACGTACTCCACTATAGCAAGTTCTCTTGGCGAGCGGTGGCTGTTCTGGGGCGGCTATGA	4273	
Db	4202	CACGTACTCCACTATTCGAAGTTCTCTTGGCGAGCGGTGGATGCTCCGGGGCGGCTATGA	4261	
Qy	4274	CATCAATAATGTGATGAGTGCACCTCAACTGACTCGACTACCATCTTTGGSCATCGGCAC	4333	
Db	4262	CATCAATAATGTGATGATGCCACTCAACTGACTCGACTACCATCTTTGGSCATCGGCAC	4321	
Qy	4334	AGTCTCGAACCAAGCGGAGACGGCTGGAGCGCGGCTCGTGTGTCGCCACCGCTACCC	4393	
Db	4322	AGTCTCGGATCAGGCAGAGACGGCTGGAGCGCGGCTCGTGTGTCGCCACCGCCACGCC	4381	
Qy	4394	TCCGGGATCGGTTACCGTGGCCACACCCCAATATCAGGAAATAGCGCTGTCCAAATGG	4453	
Db	4382	TCCGGGATCGATCAACGTTGCGACACCCCAACATCAGGAAATGGCGCTGTCCAACTGG	4441	

QY	4454	AGAGATCCCCTTCTATGTGGCAAAAGCCATCCCCATTGAGGCCATCAAGGGGGGAGGCATCT	4511
DB	4442	AGAGATTCCCTTCTATGTGGCAAAAGCCATCCCCATTGAGGCCATCAAGGGGGGAAAGGCATCT	4501
QY	4514	CATTTTCTGCGCATTTCCAAGAAAGAAATGTGACGAGCTCGCGGCAAAAGCTGACAGGCCTCGG	4573
DB	4502	CATCTTCTGCCATTTCCAAGAAAGAAAGTGTGACGAGCTCGCGGCAAAAGCTGACAGGCCTCGG	4561
QY	4574	ACTGAACGCTGTATGACATATTATACGGGGCCCTTTGATGTTCGGTCAATACCGCCTATTCGGAGA	4633
DB	4562	ACTCAATGCTGTAGCGTATTACCGGGGTCGATGTCGTCCGTATACCGACTAGCGGAGA	4621
QY	4634	CGTCGTTGTCGTGGCAACAGACGCTCTAAATGACGGGTTTACCGCGGATTTTTCGACTCAGT	4693
DB	4622	CGTCGTTGTCGTGGCAACAGACGCTCTAAATGACGGGTTTACCGCGGATTTTTCGACTCAGT	4681
QY	4694	GATCGACTGCAATACATGTGTGCCCAAGACGCGGTGTTCGCGTTCGCACAGTCGACTTCAGCTTGGATCCCACTTCAC	4753
DB	4682	GATCGACTGCAACACATGTGTGCCCAAGACGCGGTGTTCGCGTTCGCACAGTCGACTTCAGCTTGGATCCCACTTCAC	4741
QY	4754	CATTGAGACGACGACCGGTGCCCAAGACGCGGTGTTCGCGTTCGCACAGTCGAGTGAAC	4813
DB	4742	CATTGAGACGACACGCTGCCCAAGACGCGGTGTTCGCGTTCGCACAGTCGAGTAGAC	4801
QY	4814	TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCAGAGAGAACGGGCCCTCGGGCAT	4873
DB	4802	TGGCAGGGGACGAGTGGCATCTACAGGTTTGTGACTCAGAGAGAACGGGCCCTCAGGCAT	4861
QY	4874	GTTTGGATTCTTCGGTTCCTGTGTGAGTGTCTATGACGGCGGTGTGCTTGGTATGAGCTCAC	4933
DB	4862	GTTTGGACTCTTCGGTTCCTGTGTGAGTGTCTATGACGGCGGTGTGCTTGGTATGAGCTCAC	4921
QY	4934	GCCCGCTGAGACCTTCGGTTAGTTGTCGGGCTTACTAAATACACGAGGTTGCCCGTCTG	4993
DB	4922	GCCCGCTGAGACCTTCGGTTAGTTGTCGGGCTTACTAAATACACGAGGTTGCCCGTCTG	4981
QY	4994	CCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTCCAGGGCTCACCCACATAGATGCCCA	5053
DB	4982	CCAGGACCATCTAGAGTTCTGGGAGAGCGTCTTCCAGGGCTCACCCACATAGATGCCCA	5041
QY	5054	TTTCTGTGCCAGACTAAACAGGACGAGAGCACTTCTTACCTTACCTGTGGCATATCAAGC	5113
DB	5042	TTTCTGTGCCAGACCAACAGGACGAGAGCAACTCTCCCTTACCTGTGGCATATCAAGC	5101
QY	5114	TACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAAATGTGAAGTGTCT	5173
DB	5102	CACAGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGACCAAAATGTGAAGTGTCT	5161
QY	5174	CATACGGCTGAAACCTACATGCAACGGGCCAACACCCCTGCTGTATAGGCTAGAGCCGT	5233
DB	5162	CATACGGCTAAGCCCACTGTGCATGGGCCAACCGCCCTGCTGTACAGGCTAGAGCCGT	5221
QY	5234	CCAAAATGAGTTCATCTCCACACACCCCATTAATATCATATGAGCATGTCGTGGC	5293
DB	5222	TCAAAAATGAGTCACTCTCCACACACCCCATTAACCAATACATCATGGCATGTCGTGGC	5281
QY	5294	TGACCTGGAGTGTGTCATAGCACCTGGGTGCTGGTAGGCGGAGTCTTTCGAGCTTTGGC	5353
DB	5282	TGACCTGGAGTGTGTCATAGCACCTGGGTGCTGGTAGGCGGAGTCTTTCGAGCTTTGGC	5341
QY	5354	CGCATACTGCCTGACGACGAGCAGTGTGGTCTAATTGTGGGACAGGATCATCTTGTCCGGAA	5413
DB	5342	CGCGTACTGCCTGACGACGAGCAGCGTGGTCTAATTGTGGGACAGGATCATCTTGTCCGGAG	5401
QY	5414	GCCAGTGTGCTTCCCGACAGGGAAGTCTCTACACGAGGTTTCGATGAGATGGAAGATG	5473
DB	5402	GCCAGTGTATTCCCGACAGGGAAGTCTCTACACGAGGTTTCGATGAGATGGAAGATG	5461
QY	5474	TGCCTCACAACTTCTTATCATCGACAGGGAAATCAGCTCGCGAGCAATTCAGACAAA	5533
DB	5462	TGCTTTCACACCTTCTTATCATCGACAGGAATGAGCTCGCCGAGCAATTCAGACAAA	5521

Qy	5534	GGCGCTCGGGTTGTTGCAAAAGCGGCCAACAAAGAGCGGAGGCTGTGCTCTCCCTCGTGTGGA	5599
Db	5522	GGCGCTCGGATTTGTTGCAAAAGCAGCCAAAGAGCGGAGGCTGTGCTCTCCCTCGTGTGGA	5581
Qy	5594	GTCCAAGTGGGAGCCCTTCAGACCTTCTGGGCGAGACACATGTGGAATTTCATCAGCGG	5653
Db	5582	GTCCAAGTGGGAGCCCTTCAGGTCTTCTGGGCGAAACAATGTGGAATTCATCAGCGG	5641
Qy	5654	AATACAGTACTAGCAGGCTTATCCACTCTGCTCCTGGAAACCCGCGCATAGCATCATTTGAT	5713
Db	5642	GATACAGTACTTGGCAGGCCATATCCACTCTGCTCTGGAAACCCGCGATAGCATCATTTGAT	5701
Qy	5714	GGCAATTCACAGTTCATATCACTAGCCGCTCAACACCCAAAAACCCCTCTCTTTAAACAT	5773
Db	5702	GGCTTTTACAGCCTCTATCACAGCCGCTCACACCACAAATACCCCTCTCTTTAAACAT	5761
Qy	5774	CTTTGGGGGATGGGTGGCTGCCCAACTGCTCTCTCCAGCGCTGCGTCAGCTTTTCGTGGG	5833
Db	5762	CTTTGGGGGATGGGTGGCTGCCCAACTGCTCTCCCGCCAGCGCTGCTTTCCGCTTTTCGTGGG	5821
Qy	5834	CGCCGGCATCGCCGAGCGGCTGTTTGGCAGCATAGGCGTTTGGGAAGGTGCTCGTGACAT	5893
Db	5822	CGCCGGCATTTGCCGTGGGCCGTTTGGACGATAGGTCTCGGAGAGGTACTTGTGACAT	5881
Qy	5894	CTTTGGCGGCTATGGGCGAGGGTAGCCGCGCATCTGTCGGCTTTAAAGTCTATGAGCGG	5953
Db	5882	TCTGGCGGCTATGGGCGGGGTGGCTGGCGCACTCGTGCCCTTTAAAGTCTATGAGCGG	5941
Qy	5954	CGAGGTGCTTCCACCGAGAGACTGGTCAACTTACTCCCTGGCATCTCTCTCTCTGGTGC	6013
Db	5942	CGAGATGCCCTCCACTGAGGATCTGGTTAATTACTCCCTGCCATCTTTCTCTCTGGCGC	6001
Qy	6014	CCTGGTCTGGGGTCTGTGCGCAGCAATACTGCGTCGCGACGTGGGGCCCGGAGAGGG	6073
Db	6002	CCTGGTGTGGGGTCTGTGCGCAGCAATACTGCGTCGCGACGTGGGGCCCGGAGAGGG	6061
Qy	6074	GGCTGTGCAGTGAATGAACCGGCTGTAGCGTTTCGCTTCGCGGGTAAACACAGTCTCCCC	6133
Db	6062	GGCTGTGCAGTGAATGAACCGGCTGTAGCGTTTCGCTTCGCGGGTAAACACAGTCTCCCC	6121
Qy	6134	TACGCATATGTGCTTGAGAGCGACGCTGCAGACAGTGTCACTCAGATCTCTCTAGCCT	6193
Db	6122	CACGCATATGTGCTCCGAGAGCGACGCGCGGCGGTGTACTCAGATCTCTCTCCAGCCT	6181
Qy	6194	TACCATCACTCAACTGTCTGAAGCGGCTCCACCGATGGATTAAAGAGACTGCTCTAGCC	6253
Db	6182	TACCATCACTCAGTGTCTGAAGAGGCTTCATCAGTGGATTAAAGAGACTGCTCTCCAGCC	6241
Qy	6254	ATGCTCCGGCTCTGGGCTTAAGGATGTTTGGGATTTGGATATGCACGGTGTGACTGACTT	6313
Db	6242	TTGTTCCGGCTCTGGCTTAAGGATGTTTGGGACTGGATATGCACGGTGTGAGTGACTT	6301
Qy	6314	CAAGACTTGGCTCAGTCCAAATCTCTGCGCGGTTTACCGGGAGTCCCTTTCTGTCTCATG	6373
Db	6302	CAAGACTTGGCTCAGTCCAAAGCTCTCTGCGCGGTTTACCGGGACTCCCTTTCTGTCTCATG	6361
Qy	6374	CCAAAGCGGTTACAGGAGCTCTGGCGGGGGGACGGCATCATCAAAACCACTGCCCATG	6433
Db	6362	CCAAAGCGGTTACAGGAGGTCCTGGCGGGGGATGGGCATCATGCAAAACCACTGCCCATG	6421
Qy	6434	CGGAGCACAGATCGCCGGACATGTCAAAAACGGTTTCCATAGGATCGTAGGGCTAGAAC	6493
Db	6422	TGGAGCACAGATCACCGGACATGTCAAAAATGGCTCCATGAGGATTTGTTGGGCCAAAAC	6481
Qy	6494	CTGCAGCAACAGTGGCACGGAAAGTTTCCCATCAACGCATACACCGGGAAGCTTGCAC	6553
Db	6482	CTGCAGCAACAGTGGCATGGAAATTTCCCATCAACGCATACACCGGCGCCCTGCAC	6541
Qy	6554	ACCTTCCCGCGGCCCAACTATTTCAGGGCGCTATGGGGGTGCTCTGTAGGAGTACGT	6613
Db	6542	GCCCTCCCGACGCGCAACTATTTCAGGGCGCTGTGGCGGGTGGCTCTGTAGGAGTACGT	6601
Qy	6614	GGAGTTACCGGCTGTGGGGGATTTTCCACTAGCTGACCGGGATGACCACTGACAACTGAAA	6673

[illegible]

Db 7682 TCACCAAGTATGGTCTACTCCACAACATCTCCAGCGCAAGTCTCGGCGAGAAAGGT 7741
Qy 7754 CACCTTTGACAGATTCGAGTCTGGATGATCAATTACCGGGACGTACTCAAGGAGATGAA 7813
Db 7742 CACCTTTGACAGATTCGAGTCTGGAGCGACCACTACCGGGACGTGCTCAAGGAGATGAA 7801
Qy 7814 GGGAAAGGCTGCACAGTTAAGCTTAAGCTTAACTTATCTATAGAGAGGCGCTGCAAGCTGAC 7873
Db 7802 GGGAAAGGCTGCACAGTTAAGCTTAAGCTTAACTTATCTATAGAGAGGCGCTGCAAGCTGAC 7861
Qy 7874 GCGCCCACTTCGGCCAAATCCAAATTTGGCTATGGGGCAAGAGCGTCCGGAACTTATC 7933
Db 7862 GCGCCCACTTCGGCCAAATCCAAATTTGGCTATGGGGCAAGAGCGTCCGGAGCTTATC 7921
Qy 7934 CAGCAGGCGCTTAAACACATCCGCTCCGCTGTGGAGGACTTGTCTGGAAGACACTGAAC 7993
Db 7922 CAGCAGGCGCTTAAACACATCCGCTCCGCTGTGGAGGACTTGTCTGGAAGACACTGAAC 7981
Qy 7994 ACCAATTGACACCAACCATATGGCAAAAGTGAAGTCTTCTGGTCCAAACAGAGAGGG 8053
Db 7982 ACCAATTGATACCAACCATATGGCAAAAGTGAAGTCTTCTGGTCCAAACAGAGAGGG 8041
Qy 8054 AGCGCGCAAGCGAGCTCGCTTATCGTATCCAGACCTGGAGTTCGTGTATGCGAGAA 8113
Db 8042 AGCGCGCAAGCGAGCTCGCTTATCGTATCCAGACCTGGAGTTCGTGTATGCGAGAA 8101
Qy 8114 GATGGCCCTTTACAGCGTGTCTCCACCTTCCTCAGGCGGTGATGGGCTCTCATACGG 8173
Db 8102 GATGGCCCTTTACAGCGTGTCTCCACCTTCCTCAGGCGGTGATGGGCTCTCATACGG 8161
Qy 8174 ATTTCAATCTCCCGCAAGCAGCGGTCGAGTTCCTGGTGAATACCTTGAATAACAAAGAA 8233
Db 8162 ATTTCAATCTCCCGCAAGCAGCGGTCGAGTTCCTGGTGAATACCTTGAATAACAAAGAA 8221
Qy 8234 ATGCCCTATGGCTTCTCATATGACACCGCTGTTTGACTCAACCGTCACTGAGATGA 8293
Db 8222 ATGCCCTATGGCTTCTCATATGACACCGCTGTTTGACTCAACCGTCACTGAGAAATGA 8281
Qy 8294 CATTCGTGTTGAGGAGTCAATTTACCAATGTTGACTTGGCCCCGAGGCGACACAGGC 8353
Db 8282 CATTCGTGTTGAGGAGTCAATTTACCAATGTTGACTTGGCCCCGAGGCGACACAGGC 8341
Qy 8354 CATAAGTTCGCTCAGACGCGCTTTTATCATCGGGGTCCCTGTACTAACTCAAAAGGGCA 8413
Db 8342 CATAAGTTCGCTCAGACGCGCTTTTATCATCGGGGTCCCTGTACTAACTCAAAAGGGCA 8401
Qy 8414 GAATCGCGGTTATCGCGGTCGCGGCAAGTGGCGTGTGACGACTAGCTCGGTAATAC 8473
Db 8402 GAATCGCGGTTATCGCGGTCGCGGCAAGTGGCGTGTGACGACTAGCTCGGCA--- 8458
Qy 8474 CCTCATGTTACTTGAAGGCCACTGCGAGCCTGTGAGCTGCAAGCTCCAGGACTGCAC 8533
Db 8459 -----CAC 8461
Qy 8534 GATGCTCGTGAACGAGACGACCTTGTGCTTATCTGTGAAGCGCGGGAACCCAGGAGGA 8593
Db 8462 GATGCTCGTGAACGAGACGACCTTGTGCTTATCTGTGAGAGTGGGGAACCCAGGAGGA 8521
Qy 8594 TGCGCGGCGCTTACGAGCGCTTACGAGGCTATGACTAGTATTCGCGCCCCCGCGGGA 8653
Db 8522 TGCGCGGCGCTTACGAGCGCTTACGAGGCTATGACTAGTATTCGCGCCCCCGCGGGA 8581
Qy 8654 TCGCGCCCAACAGAAATAGGACTGGAGTGAATACATCATGTTCTCTCAATGTGTGAGT 8713
Db 8582 CCGCGCCCAACAGAAATAGGACTTGGAGCTGATAACGTCATGCTCCTCAATGTGTGAGT 8641
Qy 8714 CGCGCAGCATGCTCGGCAAAAGGTTACTTACTCTACCGGTGACCCACACCCCGCT 8773
Db 8642 CGCGCAGCATGCTCGGCAAAAGGTTACTTACTCTACCGGTGACCCACACCCCGCT 8701
Qy 8774 TGACGCGGCTGCTGGGAGACAGCTAGACACACTCCAATCACTTTGGCTAGGCAATAT 8833
Db 8702 CGCACGGGCTGGTGGGAGACAGTTAGACACACTCCAGTCAACTCTTGGCTAGGCAATAT 8761

Qy 8834 CATCATGTATGCGCCACCCCTATGCGCAAGGATGATTCTGTGAGACTCAGCTTTTCTCCAT 8893
Db 8762 CATCATGTATGCGCCACCCCTATGCGCGAGGATGATTCTGTGAGACTCAGTTTCTCTAT 8821
Qy 8894 CTTCTAGCTCAAGAGCAACTTGAAGAAAGCCCTGGATTGTTCAGATCTACGGGCTTGCTA 8953
Db 8822 CTTCTAGCTCAGGAGCAACTTGAAGAAAGCCCTGGATTGTTCAGATCTACGGGCTTGCTA 8881
Qy 8954 CTCATTGAGCCACTTGAACCTCAGATCATTAAGCAAGCTCCATGCTTACGCGCAAT 9013
Db 8882 CTCATTGAGCCACTTGAACCTCAGATCATTAAGCAAGCTCCATGCTTACGCGCAAT 8941
Qy 9014 TACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGTGGCTTTCATGCTCAGGAACT 9073
Db 8942 TTCACTCCACAGTTACTCTCCAGGTGAGATCAATAGGTGGCTTTCATGCTCAGGAACT 9001
Qy 9074 TGGGTACCACTTCGAACTTGGAGACATTCGGGCGCAGAAAGTGTCCGCGCTAAGCTACT 9133
Db 9002 TGGGTACCGCTTTCGGAGTCTGGAGACATTCGGGCGCAGAAAGTGTCCGCGCTAAGCTACT 9061
Qy 9134 GTCCAGGGGGGAGGCGGCCACTTGTGGCAGATACCTCTTTAACTGGGCGAGTAAGGAC 9193
Db 9062 GTCCAGGGGGGAGGCGGCCACTTGTGGCAGATACCTCTTTAACTGGGCGAGTAAGGAC 9121
Qy 9194 CAAGCTTAACTCACTTCAATCCCGCTGCTCCAGCTAGACTTGTCCGCGCTGCTGCT 9253
Db 9122 CAAGCTTAACTCACTTCAATCCCGCTGCTCCAGCTAGACTTGTCCGCGCTGCTGCT 9181
Qy 9254 CGCTGGTTACAGCGGGGAGACATATATACAGCTGCTCGTCCCGCAGACCCCGCTGCT 9313
Db 9182 TGCTGGTTACAGCGGGGAGACATATATACAGCTGCTCGTCCCGCAGACCCCGCTGCT 9241
Qy 9314 TCCTGTGCTACTCTCTACTTTCTGTAGGGTAGGCAATTAAGCTTCCCTGCTGCTGCT 9373
Db 9242 CATGTTGCTACTCTCTACTTTCTGTAGGGTAGGCAATTAAGCTTCCCTGCTGCTGCT 9301
Qy 9374 AACGGGAGCTTACCACTCAGGCTTAAAGCAATTTCTGTTTCTGTTTCTGTTTCT 9424
Db 9302 AACGGGAGCTTAACTCAGGCTTAAAGCAATTTCTGTTTCTGTTTCTGTTTCTGTTT 9352

RESULT 5
US-10-333-449A-33
; Sequence 33, Application US/10333449A
; Publication No. US20040137424A1
; GENERAL INFORMATION:
; APPLICANT: Tan, Yin Hwee
; APPLICANT: Lim, Siew Pheng
; APPLICANT: Lim, Seng Gee
; APPLICANT: Hong, Wan Jin
; TITLE OF INVENTION: NUCLEIC ACIDS AND METHODS FOR DETECTING VIRAL INFECTION,
; UNCOVERING ANTI-VIRAL DRUG CANDIDATES AND DETERMINING DRUG
; TITLE OF INVENTION: RESISTANCE OF VIRAL ISOLATES
; FILE REFERENCE: 01/22137
; CURRENT APPLICATION NUMBER: US/10/333.449A
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 9609
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-10-333-449A-33

Query Match 84.1%; Score 8067; DB 7; Length 9609;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 8655; Conservative 0; Mismatches 940; Indels 14; Gaps 1;
Qy 1 GCCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAACTACTG 60
Db 1 GCCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAACTACTG 60

Qy	61	TC	TT	C	A	C	G	A	A	A	G	C	G	T	C	T	A	G	C	A	T	G	C	T	T	A	G	T	A	G	T	C	T	G	C	A	G	C	T	C	C	A	G	A	C	120							
Db	61	TC	TT	C	A	C	G	C	A	A	A	G	C	G	T	C	T	A	G	C	A	T	G	C	T	T	A	G	T	A	G	T	C	T	G	C	A	G	C	T	C	C	G	G	A	C	120						
Qy	121	C	C	C	C	T	C	C	G	G	A	G	A	G	C	C	A	T	A	G	T	C	T	C	G	G	A	C	G	G	T	A	G	T	A	G	T	C	A	C	C	G	A	A	T	C	C	A	G	180			
Db	121	C	C	C	C	T	C	C	G	G	A	G	A	G	C	C	A	T	A	G	T	C	T	C	G	G	A	C	G	G	T	A	G	T	A	G	T	C	A	C	C	G	A	A	T	C	C	A	G	180			
Qy	181	G	A	C	A	C	C	G	G	T	C	T	T	T	C	T	T	G	A	T	C	A	A	C	C	C	C	T	C	A	A	T	G	C	C	T	C	A	A	T	G	C	C	T	G	C	C	C	C	240			
Db	181	G	A	C	A	C	C	G	G	T	C	T	T	T	C	T	T	G	A	T	C	A	A	C	C	C	C	T	C	A	A	T	G	C	C	T	C	A	A	T	G	C	C	T	G	C	C	C	C	240			
Qy	241	G	C	G	A	C	T	G	T	A	G	C	C	G	A	T	A	G	T	T	T	G	G	T	C	G	C	A	A	A	G	G	C	T	T	G	T	A	G	T	A	G	C	T	G	A	T	A	G	300			
Db	241	G	C	G	A	C	T	A	G	C	C	G	A	T	A	G	T	T	T	G	G	T	C	G	C	A	A	A	G	G	C	T	T	G	T	A	G	T	A	G	C	T	G	A	T	A	G	300					
Qy	301	G	T	G	T	T	G	C	G	A	G	T	C	T	C	T	A	G	C	C	G	A	G	T	C	T	G	T	A	G	C	C	A	T	A	G	C	C	A	T	A	G	C	A	G	A	A	T	C	C	A	G	360
Db	301	G	T	G	T	T	G	C	A	G	T	C	T	C	T	A	G	C	C	G	A	G	T	C	T	G	T	A	G	C	C	A	T	A	G	C	C	A	T	A	G	C	A	G	A	A	T	C	C	A	G	360	
Qy	361	C	T	C	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	420				
Db	361	C	T	C	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	420					
Qy	421	G	T	G	T	C	A	G	A	T	C	G	T	T	G	T	A	G	C	T	T	T	G	C	G	C	A	G	C	A	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	480			
Db	421	G	T	G	C	C	A	G	A	T	C	G	T	T	G	T	A	G	C	T	T	T	G	C	G	C	A	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	480				
Qy	481	G	C	G	C	G	A	C	T	A	G	A	G	C	T	C	C	A	G	C	G	H	T	C	G	A	A	C	C	T	C	G	T	G	A	A	G	G	C	A	A	C	C	T	A	T	C	C	C	A	A	540	
Db	481	G	C	G	C	A	T	A	G	A	G	A	C	T	C	C	A	G	C	G	H	T	C</																														

[illegible]

Db 2221 TCAAGGTTAGGATGTATGTGGGGGGCGTAGAGCACAGGCTCAACGGTCACGGCACTGGA 2280
 Qy 2281 CTCAGGAGACGCTGTAACTTGGAGAGACAGGATAGGTGAGAACTCAGCCGCTGCTGC 2340
 Db 2281 CTCAGGAGACGCTGTGACTTGGAGAGACAGGATAGATCAGAGCTTAGCCGCTGCTGC 2340
 Qy 2341 TGTCTACACAGAGTGGCAGATACTGCCCTGTGCTTTTCCACCCTACCGGCTTTATCCA 2400
 Db 2341 TGTCTACGACAGAGTGGCAGATACTGCCCTGTGCTTTTCCACCCTACCGGCTCTGTCCA 2400
 Qy 2401 CTGGTTTGTATCCATCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGGTAGGGT 2460
 Db 2401 CTGGTTTGTATCCATCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGGTAGGGT 2460
 Qy 2461 CAGGTTTGTCTCTCTTTGCAATCAAAATGGAGTACATCCCTGTGCTTTTCTCTCTCTGG 2520
 Db 2461 CAGTGGTTGTCTCCCTCGTGATCAAAATGGGAGTATGCTTGTCTCTCTCTCTCTGG 2520
 Qy 2521 CAGACGGCGGCTGTGCTCTGTCTGTGATGATGCTGTGCTGTGATGATGATGATGATGATG 2580
 Db 2521 CAGACGGCGGCTGTGCTCTGTCTGTGATGATGCTGTGCTGTGATGATGATGATGATGATG 2580
 Qy 2581 CCTTAGAGACTTGGTGGTCTCAATCGCGGCTCCGTGGCGGAGCGCATGATGATGATGATG 2640
 Db 2581 CCTTAGAGAACTTGGTGGTCTCAATCGCGGCTCCGTGGCGGAGCGCATGATGATGATGATG 2640
 Qy 2641 CCTTCTGTGTCTCTCTGGCGGCTGTGATCAATTAAGGGCAGGCTGGCTCTCTGGGCGG 2700
 Db 2641 CCTTCTGTGTCTCTCTGGCGGCTGTGATCAATTAAGGGCAGGCTGGTCTCTCTGGGCGG 2700
 Qy 2701 CGTATGCTTTTATGGCGGTATGGCGCTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 Db 2701 CATATGCTTTTATGGCGGTATGGCGCTGTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 Qy 2761 CTTAGGCTTTGGACCGGAGATGCTCATCGTGGGGGCTGGCTCTCTCTCTCTCTCTCTCTCT 2820
 Db 2761 CGTACGCTATGGACCGGAGATGCTCATCGTGGGGGCTGGCTCTCTCTCTCTCTCTCTCTCT 2820
 Qy 2821 TATCTTTGACCTGTACCACTACAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2880
 Db 2821 TACTCTTGACCTGTACCACTATAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2880
 Qy 2881 AATACTTTATCACAGAGCGAGCGGACATGCAAGTGTGGGTCTCTCTCTCTCTCTCTCTCTCTCT 2940
 Db 2881 AGTACTTTATCACAGAGCGAGGCAATTTTGCACGTGTGGGTCTCTCTCTCTCTCTCTCTCTCTCT 2940
 Qy 2941 GGGAGGCGCGATGCCATCATCTCTCTCAGTGTGGGTTCATCCAGAGTAAATTTTGTG 3000
 Db 2941 GAGGGGCGGATGCCGTCTCTCTCTCAGTGTGGGTTCATCCAGAGTAAATTTTGTG 3000
 Qy 3001 ACATCACCAACTCTGTCTGCCATCTCGGCCCGCTCATGCTGTCTCGAGGCTGGCATAA 3060
 Db 3001 ACATCACCAACTCTGTCTGCCATCTCGGCCCGCTCATGCTGTCTCGAGGCTGGCATAA 3060
 Qy 3061 CGAGAGTCCGCTACTCTGTGGCGCTCAAGGGCTCATTCGTCATGATGATGTAGTGGAA 3120
 Db 3061 CAAGGTGCCATCTCTGTGGCGGCCCAAGGCTCATTCGTCGCTGGCGGTTAGCGGGA 3120
 Qy 3121 AAGTCCCGGGGCTCATTTATGTCMAATGGTCTTCATGAAGCTGGCGGCTGACAGTA 3180
 Db 3121 AAGTCCCGGGGCTCATTTATGTCMAATGGTCTTCATGAAGCTGGCGGCTGACAGTA 3180
 Qy 3181 AGGTCCGCTGGGGGCTCATTTATGTCMAATGGTCTTCATGAAGCTGGCGGCTGACAGTA 3240
 Db 3181 AGGTCCGCTGGGGGCTCATTTATGTCMAATGGTCTTCATGAAGCTGGCGGCTGACAGTA 3240
 Qy 3241 TTGCGGCTGGGCTAGAGCGGCTGTCTCTCTCGCCATGGAGACCAAGGTCTATCATCTGGG 3300
 Db 3241 TTGCGGCTAGAGCTGAGCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3300
 Qy 3301 GAGCAGACACCGCTGTGGGACATCATCTTTGGGCTCTACCGGTCTCGGCCCGGAGGG 3360
 Db 3301 GGGCAGACACCGCTGTGGGACATCATCTTTGGGCTCTACCGGTCTCGGCCCGGAGGG 3360

Qy 3361 GGAAAGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTTGGCGACTCTCTTG 3420
 Db 3361 GGAGGAGATATTTTGGGACCGGCTGATAGTCTTGAAGGGCAGGGTGGCGACTCTCTTG 3420
 Qy 3421 CGCCCATCACGGCTACTCTCCAAACAAACCGGGGCGTACTTTGGTTGCATCATCACTAGCC 3480
 Db 3421 CGCTATTAGGCTACTCTCCAGCAGACGGGGGCTACTTTGGCTGCATCATCACTAGTC 3480
 Qy 3481 TCACAGCCGGGCAAGAAACCCAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
 Db 3481 TCACAGCTGGGCAAGAAACCCAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
 Qy 3541 AATCTTTCTGGGACCTCATCAACGGGCTGTGCTGGAAGTCTACCATGGGCTGGCT 3600
 Db 3541 AGTCTTTCTAGGACCTCGCTCAACGGGCTGTGTTGGAGTGTCTTCCATGGTGGCT 3600
 Qy 3601 CGAAGACCTTAGCGGCTCAAAAGGTCCAATCAACCCAAATGTACCAATGTAGACCTGG 3660
 Db 3601 CAAAAACCTTAGCGGCTCAAAAGGGGCAATTAATCAATGTATACCAACGTGACAGG 3660
 Qy 3661 ACCTCGTGGCTGGCAGGCGCCCCCGGGGCGCTCCATGACACCATGACGTGTGGCA 3720
 Db 3661 ACCTCGTGGCTGGCAGGCGCCCCCGGGGCGCTCCATGACACCATGACCTGTGGCA 3720
 Qy 3721 GCTCGGACCTTTACTTTGGTCAAGACATGCTGATGCTCATTTCCGGTGGCGGGAGCG 3780
 Db 3721 GCTCGGACCTTTACTTTGGTCAAGACATGCTGATGCTCATTTCCGGTGGCGGGAGCG 3780
 Qy 3781 ACAGCAGGGAAGTCTACTCTCTCCAGGCGCTCTCTACTCTGAAAGGCTCTCTCGGGTG 3840
 Db 3781 ACACAGAGGAGGCTACTCTCTCCAGGCGCTCTCTACTTAAAGGCTCTCTCGGGTG 3840
 Qy 3841 GTCCATTTGCTTCCCTTCCGGGACAGTCTGTGGCGGCTCTTCCGGGCTGTGTGTGACCC 3900
 Db 3841 GTCCACTCTCTCGGCTGGGACATGCTGTAGGCACTCTTCCGGGCGCGGTATGACCC 3900
 Qy 3901 GGGGGTGGGAGGCGGTGACTTATACCGTGTAGTCTATGGAGACTACCAATGGGT 3960
 Db 3901 GGGGGTGGGAGGCGGTGACTTATACCGTGTAGTCTATGGAGACTACCAATGGGT 3960
 Qy 3961 CTCCGCTCTCACAGACAACTCAACCCCGGCTGTACCGCAGACATTCCAAGTGGCAC 4020
 Db 3961 CCGGCTCTCACAGACAACTCAACCCCGGCTGTACCGCAGACATTCCAAGTGGCTC 4020
 Qy 4021 ATCTGACGCTCTCTACTTGGCAGCGGCAAGAGCAACAAAGTCCCGGCTGCGTATGACGCC 4080
 Db 4021 ATCTGACGCTCTCTACTTGGCAGCGGCAAGAGCAACAAAGTCCCGGCTGCGTATGACGCC 4080
 Qy 4081 AAGGTTACAGGCTCTCTCTGAAACCGGCTGGTGGCGGCACTTAAAGTGGGTTGGGGGCT 4140
 Db 4081 AAGGTTACAGGTTCTCTCTGAAACCGGCTGGTGGCGGCACTTAAAGTGGGTTGGGGGCT 4140
 Qy 4141 ATATGTCAGGACACGGTATCGACCTTAACATCAGACCTGGGTTAAGGACCATTAACA 4200
 Db 4141 ATATGTCAGGACATATGGTGTGATCTTAACATCAGACCTGGGTTAAGGACCATTAACA 4200
 Qy 4201 CGGGCGCTCCATTAACGCTTCTCACTATGGAAGTCTCTTGGCGAGCGGTGGCTGTTGTG 4260
 Db 4201 CGGGCGCTCCATTAACGCTTCTCACTATGGAAGTCTCTTGGCGAGCGGTGGCTGTTGTG 4260
 Qy 4261 GGGCGCTTATGACATCATTAATATGATGATGCCACTCAACTGACTCGATACCATCT 4320
 Db 4261 GGGGAGCTTACGACATCATTAATATGATGATGCCACTCAACTGACTCGATACCATCT 4320
 Qy 4321 TGGGCATCGGACAGTCTTGGACCAAGGGAGCGGTGGAGCGGCTCGCTGCTGCTCG 4380
 Db 4321 TGGGCATCGGACAGTCTTGGACCAAGGGAGCGGTGGAGCGGCTCGCTGCTGCTCG 4380
 Qy 4381 CCACGCTTACACCTCCGGGATCGGTTACCGGTGCCACACCCCAATATCAGGAAATAGGCC 4440
 Db 4381 CCACGCTTACACCTCCGGGATCGGTTACCGGTGCCACATCCCAATATCAGGAGATCGCTC 4440

Qy	4441	TGTC	CAACA	CAATG	GAGAGAT	CCCC	TTCTAT	TGG	CAAA	AGCCAT	CCCC	ATATG	AGGCCAT	CAAGG	4500		
Db																	
4441	TAT	CAACA	CACTG	GAGAGAT	CCCC	TTCTAT	TGG	CAAA	AGCCAT	CCCC	ATATG	AGGCCAT	CAAGG	4500			
Qy	4501	GGGG	AGGCAT	CTCAT	TTTTCTG	CCAT	TCC	CAAG	AGAA	TGT	GAC	GAGCT	CGCG	CAAGC	4560		
Db																	
4501	GGGG	AGGCAT	CTCAT	TTTTCTG	CCAT	TCC	CAAG	AGAA	TGT	GAC	GAGCT	CGCTG	CAAGT	4560			
Qy	4561	TGAC	AGCC	CTCG	GA	CTGAA	CGCTGT	AG	CA	TAT	TAC	CGGG	CCCTG	ATGT	CGTCATAC	4620	
Db																	
4561	TGT	CGGCC	CTCG	GA	CTGTAA	CGCTGT	AG	CA	TAT	TAC	CGGG	CCCTG	ATGT	CGTCATAC	4620		
Qy	4621	CGC	TAT	CG	GAG	CGT	CTTGT	CG	TG	CA	A	CGCT	CTA	TG	ACGGGTTT	4680	
Db																	
4621	CAAC	TAG	CG	GAG	CGT	CTTGT	CG	TG	CA	A	CGCT	CTA	TG	ACGGGTTT	4680		
Qy	4681	ATTT	TG	ACT	CAG	TGAT	CG	AT	CA	TAT	GTG	T	CA	CCG	AG	4740	
Db																	
4681	ATT	CCG	ACT	CAG	TGAT	CG	AT	CA	TAT	GTG	T	CA	CCG	AG	4740		
Qy	4741	ATCC	CA	CTT	CAC	CAT	TG	AG	CA	CG	CGT	GC	CC	CA	AG	4800	
Db																	
4741	ACC	T	ACT	T	CAC	CAT	TG	AG	CA	CG	CGT	GC	CC	CA	AG	4800	
Qy	4801	GGC	GAG	T	AG	AA	CTG	G	AG	G	G	T	T	G	A	4860	
Db																	
4801	GGC	GAG	C	AG	G	ACT	T	G	G	G	T	T	G	A	4860		
Qy	4861	GGCC	CT	CG	GC	AT	TTCT	CG	GT	CT	GT	G	AG	T	G	4920	
Db																	
4861	GGCC	CT	CG	GC	AT	TTCT	CG	GT	CT	GT	G	AG	T	G	4920		
Qy	4921	GGT	AT	G	AG	CT	C	AG	CC	T	CG	GT	T	AG	TTG	4980	
Db																	
4921	GGT	AT	G	AG	CT	C	AG	CC	T	CG	GT	T	AG	TTG	4980		
Qy	4981	GTT	T	CC	CG	CT	CG	CA	GG	AC	CA	CT	G	AG	AG	5040	
Db																	
4981	GTT	T	CC	CG	CT	CG	CA	GG	AC	CA	CT	G	AG	AG	5040		
Qy	5041	ACAT	AG	T	AG	TC	CC	AG	ACT	TAA	AC	AG	G	AG	G	5100	
Db																	
5041	ACAT	AG	CC	CA	CTT	CTT	G	T	G	T	C	T	CA	G	A	5100	
Qy	5101	TGG	CAT	AT	CA	G	CTT	CA	G	T	G	CG	CG	CT	CA	AG	5160
Db																	
5101	TAG	CGT	AC	AG	CTT	CA	G	T	G	CG	CG	CG	CG	AG	G	5160	
Qy	5161	TGT	GGA	G	T	G	T	C	A	T	C	CG	G	T	G	A	5220
Db																	
5161	TGT	GGA	G	T	C	T	C	A	T	C	CG	G	T	G	A	5220	
Qy	5221	GGC	T	AG	G	AG	CC	G	T	CA	C	CA	CC	CC	AT	AA	5280
Db																	
5221	GGC	T	AG	G	AG	CC	G	T	CA	C	CA	CC	CC	AT	AA	5280	
Qy	5281	CAT	G	CAT	G	T	CG	G	AG	T	CG	T	CAT	AG	CA	5340	
Db																	
5281	CAT	G	CAT	G	T	CG	G	AG	T	CG	T	CAT	AG	CA	5340		
Qy	5341	TTG	CAG	CTT	T	G	CG	C	T	A	C	T	G	A	G	5400	
Db																	
5341	TG	C	A	G	CC	CT	T	G	A	C	A	G	CC	CT	T	5400	
Qy	5401	TCT	T	T	CG	G	A	G	AG	C	C	A	G	AG	T	5460	
Db																	
5401	TCT	T	T	CG	G	A	G	AG	C	C	A	G	AG	T	5460		
Qy	5461	AG	T	G	A	G	A	G	AG	T	CG	C	T	C	A	5520	
Db																	
5461	AG	T	G	A	G	A	G	AG	T	CG	C	T	C	A	5520		
Qy	5521	AAT	T	CA	AG	CA	AA	AG	CG	CT	CG	GG	T	GT	G	5580	
Db																	
5521	AAT	T	CA	AG	CA	AA	AG	CG	CT	CG	GG	T	GT	G	5580		

[illegible]


```
QY 8821 GGCTAGGCAATATCATCATATGATGCGCCACCCCTATGGGCAAGATGATTTCTGATGACTC 8880
Db 8821 GGCTAGGCAATATCATCATATGATGCGCCACCCCTATGGGCAAGATGATTTCTGATGACTC 8880
QY 8881 ACTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTGAAGAGCCCTGGATTGTGAGATCT 8940
Db 8881 ACTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTGAAGAGCCCTGGATTGTGAGATCT 8940
QY 8941 ACGGGGCTTGCTACTCCATCTGAGCCACTTGACCTACCTCAGATCATTTGAAGCACTCCATG 9000
Db 8941 ACGGGGCTTGCTACTCCATCTGAGCCACTTGACCTACCTCAGATCATTTGAAGCACTCCAG 9000
QY 9001 GTCCTAGGCAATTTACATCCACAGTTACTCTCCAGGTGAGATCAATAGGTGGCTTTCAT 9060
Db 9001 GTCCTAGGCAATTTTCGCTCCATAGTTACTCTCCAGGCGAGATCAATAGGTGGCTTTCAT 9060
QY 9061 GCCTCAGGAACTTGGGTACCACTTCCGACCTTGGCACTGAGAGATCGGCGCAGAGTGCC 9120
Db 9061 GCCTCAGGAACTTGGGTACCACTTCCGACCTTGGCACTGAGAGATCGGCGCAGAGTGCC 9120
QY 9121 GGGCTAAGCTACTGCTCCAGGGGGGAGGGCGCCACTTGTGGCAGATACCTCTTTAACT 9180
Db 9121 GGGCTAAGCTACTGCTCCAGGGGGGAGGGCGCCACTTGTGGCAGATACCTCTTTAACT 9180
QY 9181 GGGCAGTAAGGACCAAGCTTAACTCACTCCAAATCCCGGCGCGCTCCAGCTGGACTTGT 9240
Db 9181 GGGCAGTAAGGACCAAGCTTAACTCACTCCAAATCCCGGCGCGCTCCAGCTGGACTTGT 9240
QY 9241 CTGGCTGGTTTCTGCTGTTACAGGGGGGAGACATATATACAGCTGCTGCTGGCC 9300
Db 9241 CTGGCTGGTTTCTGCTGTTACAGGGGGGAGACATATATACAGCTGCTGCTGGCC 9300
QY 9301 GACCCCGCTGCTTCCGTTGCTGCTTACTTCTACTTCTGTTAGGGGTAGGCATTTACCTGC 9360
Db 9301 GACCCCGCTGCTTCCGTTGCTGCTTACTTCTACTTCTGTTAGGGGTAGGCATTTACCTGC 9360
QY 9361 TCCCAACCGATGAACGGGGAGCTAACCACTCCAGGCGCTTAAGCCATTTCTCTGTTTTT 9420
Db 9361 TCCCAACCGATGAACGGGGAGCTAACCACTCCAGGCGCTTAAGCCATTTCTCTGTTTTT 9420
QY 9421 TTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9480
Db 9421 TTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9480
QY 9481 TTTTCTCCCTCTT-----TAATGCTGCTCCATCTTAGCCCTAGTCACGCG 9526
Db 9481 TTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCTTTTGGTCCATCTTAGCCCTAGTCACGCG 9540
QY 9527 TAGCTGTGAAGGTCGCTGAGCGCATGACTGACAGAGTGTCTGATCTGGCCTCTCTGC 9586
Db 9541 TAGCTGTGAAGGTCGCTGAGCGCATGACTGACAGAGTGTCTGATCTGGCCTCTCTGC 9600
QY 9587 AGATCATGT 9595
Db 9601 AGATCATGT 9609
```

RESULT 6

```
US-10-259-275-39
; Sequence 39, Application US/10259275
; Publication No. US20030125541A1
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0120
; CURRENT APPLICATION NUMBER: US/10/259,275
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/747,419
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/325,236
```

```
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 9275
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding hepatitis C virus polyprotein
; OTHER INFORMATION: derived from HCV-N
US-10-259-275-39
```

Query Match

Best Local Similarity 82.2%; Score 7885.8; DB 6; Length 9275;

Matches 8437; Conservative 0; Mismatches 817; Indels 21; Gaps 4;

```
QY 342 ATGAGCAGCAATCTCTAAACCTCAAAAGAAAAACCAAAACGTAACACCAACCGCGCCACAG 401
Db 1 ATGAGCAGCAATCTCTAAACCTCAAAAGAAAAACCAAAACGTAACACCAACCGCGCCACAG 60
QY 402 GACCTCAAGTTCCCGGGGGTGGTCCAGATCGTTGGTGGAGTTTACCTGTCGCGCGCAGG 461
Db 61 GACCTCAAGTTCCCGGGGGTGGTCCAGATCGTTGGTGGAGTTTACCTGTCGCGCGCAGG 120
QY 462 GCGCCAGGTTGGGTGTGCGCGCGACTAGGAAGGCTTCCGAGCGGTCCGAACCTCGTGA 521
Db 121 GCGCCAGGTTGGGTGTGCGCGCGCATCAGGAAGACTTCCGAGCGGTCCGAACCTCGTGA 180
QY 522 AGGCGCAACCACTATCCAAAGGCTGCGCGACCCGAGGCGAGGGCTTGGGCTCAGCCCGG 581
Db 181 AGGCGCAGACCTATCCCAAGGCTCGCGCGCCGAGGCGAGGGCTTGGGCTCAGCCCGG 240
QY 582 TACCTTGGCCCTCTATAGGCAATGAGGGCTTGGGGTGGGAGGATGGCTCTGTGACCC 641
Db 241 TATCTTGGCCCTCTATAGGCAATGAGGGCTTGGGGTGGGAGGATGGCTCTGTGACCC 300
QY 642 CGCGCTCCCGGCTAGTTGGGGCCACGAGACCCCGCGGTAGGTCCGCTAACTTGGGT 701
Db 301 CGCGCTCCCGGCTAGTTGGGGCCACGAGACCCCGCGGTAGGTCCGCTAACTTGGGT 360
QY 702 AAGGTCATGATACCTTACATGCGGGTTCGCGGATCTCATGGGTAATTCGCTCGTC 761
Db 361 AAGGTCATGATACCTTACATGCGGGTTCGCGGATCTCATGGGTAATTCGCTCGTC 420
QY 762 GCGCGCCCTTAGGGGGGCTGCGAGGGCTTGGCAGCGGTGTCGGGTCTGGAGGAC 821
Db 421 GCGCGCCCTTAGGGGGGCTGCGAGGGCTTGGCAGCGGTGTCGGGTCTGGAGGAC 480
QY 822 GCGGTGAATATGCAACAGGGAACCTTGGCCGGTGTCTTCTCTATCTTCTCTTGGCT 881
Db 481 GCGGTGAATATGCAACAGGGAACCTTGGCCGGTGTCTTCTCTATCTTCTCTTGGCT 540
QY 882 CTGCTGTCTCTTTGACCATCCAGCTTCCGCTTATGAAGTCCGCAACGCTGTCGGGATA 941
Db 541 CTGCTGTCTCTTTGACCATCCAGCTTCCGCTTATGAAGTCCGCTTATGAAGTCCGCGGTA 600
QY 942 TACCATGTGCAAGCACTGCTCCAACTCAAGCATTTGTATGAGGAGCGAGCATGATC 1001
Db 601 TACCATGTGCAAGCACTGCTCCAACTCAAGCATTTGTATGAGGAGCGAGCATGATC 660
QY 1002 ATGCATATCCCGGGTGGTGGCTGTGTTGAGGAGGTAAACAGCTCCGTTGCTGGGTA 1061
Db 661 ATGCATATCCCGGGTGGTGGCTGTGTTGAGGAGGTAACTCTTCCGCTGCTGGGTA 720
QY 1062 GCGCTCACTCCACAGCTCGCGGCCAGGAATGCAAGCTCCCACTCCCACTACGACAC 1121
Db 721 GCGCTCACTCCACAGCTCGCGGCCAGGAATGCAAGCTCCCACTACGACAC 780
QY 1122 CACGTGCACTTGTCTGTTGGAGCGGCTGTTTCTGCTCGCTATGTAGTGGGGATCTC 1181
Db 781 CACGTGCACTTGTCTGTTGGAGCGGCTGTTTCTGCTCGCTATGTAGTGGGGATCTC 840
```

1182 TCGGATCTATTTTCCTCGTCCAGCTGTTCACCTTCTCGCCTCGCGGATGAGCA 1241
 Db |||||
 841 TCGGATCTGTTTCTCTGCTCTCACTAGCTTCACTTCTCGCCCCCGCGGATGCGCA 900
 Qy |||||
 1242 GTGAGGACTGCAACTGCTCAATCTATCCGGCCATGATCAGGTCAACGCGATGGCTCG 1301
 Db |||||
 901 TTGAGGACTGCAATTTGTCGATCTACCCCGGCGACGCTCAGGTCAACGCGATGGCTCG 960
 Qy |||||
 1302 GATATGATGATGAATGCTGCTCACTCAACAGAGCCCTAGTGTGTCGAGTTGCTCGGATC 1361
 Db |||||
 961 GACATGATGATGAATGCTGCTCACTCAACAGAGCCCTAGTGTGTCGAGTTGCTCGGATC 1020
 Qy |||||
 1362 CCAAGAGCTGCTGAGATGATGTCGCGGGGGGCCCACTGGGGAGTCTCGGGGGCTTGGC 1421
 Db |||||
 1021 CCACAGAGCTGCTGAGATGATGTCGCGGGGGGCCCACTGGGGAGTCTCGGGGGCTTGGC 1080
 Qy |||||
 1422 TACTATTCATGATGAGGACTGGCTAGGTTCTGATGTCGCTACTCTTTGCGGC 1481
 Db |||||
 1081 TACTATTCATGCGGGGAACTGGGCTAAGGTTTGAATGATGCTACTTTTTCGCGC 1140
 Qy |||||
 1482 GTTGAGGGGAGACCCACAGAGCGGAGGCTGGCGGCCACACCACTCCCGGTTCAAG 1541
 Db |||||
 1141 GTTGAGGGGACACCTCTCAACAGGGGGGCGCTGCCGCCCTCACAGCGGTTTCGCG 1200
 Qy |||||
 1542 TCCCTTTTCTCATCTGGGGGCTCTAGAAATCCAGCTTGTGAATACCAACGCGAGTGG 1601
 Db |||||
 1201 GGCTCTTTACCTGGCGGCTCTCAGAGAACTCAGCTTATAAACACCAATGGCAGTTGG 1260
 Qy |||||
 1602 CACATCAACAGGACTGGCTTAATTCGATGATGCTCCCTCAAACTGGGTTCTTTGCGCG 1661
 Db |||||
 1261 CACATCAACAGGACTGGCTTAATTCGATGATGCTCCCTCAGAGCTGGGTTCTTTGCGCG 1320
 Qy |||||
 1662 CTGTTTACGACACAAAGTTCAACTCGTCCGGGTGCGCGAGCGCATGGCGAGCTGCGCG 1721
 Db |||||
 1321 CTGTTTACGACACAAAGTTCAACTCGTCCGGGTGCGCGAGCGCATGGCGAGCTGCGCG 1380
 Qy |||||
 1722 CCAATTCAGTGTTCGCCAGGGTGGGCCCATCACTATCACTAAGCT---AACAGC 1778
 Db |||||
 1381 TCCATTGACAAAGTTTCAGCAGGAGTGGGTCCTATCACTTATGCTGAGCCTACAAAAGAC 1440
 Qy |||||
 1779 TCGGATCAGAGGCTTATTCGTGCGCATACGCGCTCGACCGTGTGCTGTCGACCGCG 1838
 Db |||||
 1441 CCGGACAGAGGCTTATTCGTGCGCATACCGCTCAACAAATGTGGTATCGTACCTGGC 1500
 Qy |||||
 1839 TCCAGGTGTGTCGAGGATTTGTTTCAACCCAGCCCTGTTGTTGGTGGGACCAAC 1898
 Db |||||
 1501 TCGAGGTGTGTCGAGGATTTGCTTCAACCCAGCTCTGTTGTCGTTGGGAGCAAC 1560
 Qy |||||
 1899 GATCGTTCCGGTGTCCCTAGCTATAGCTGGGGGAGAAATGAGACAGAGTGTGCTCCTC 1958
 Db |||||
 1561 GATCGTCTCGCAACCTACGTACAGCTGGGGGAGAAAGATGACGCTGCTCCTT 1620
 Qy |||||
 1959 AACAAAGGCTCGCGCACAGGCACTGTTTGGCTGTATCAGGATGAATAGTACTGGG 2018
 Db |||||
 1621 AACAAAGGCTCGCGCACAGGCACTGTTTGGCTGTATCAGTGAATAGTACTGGG 1680
 Qy |||||
 2019 TTCACTAAGAGCTGGGAGGTCCTCGCTGAACATCGGGGGGTCGGTAAACCGCACCTTG 2078
 Db |||||
 1681 TTCACTAAGAGCTGGGAGGTCCTCGCTGAACATCGGGGGGTCGGTAAACCGCACCTTG 1740
 Qy |||||
 2079 ATCTGCCCCAGGACTGCTTCCGGAAGCACCCCGAGGCTACTTACACAAAATGTGGCTCG 2138
 Db |||||
 1741 ACCTGCCCCAGGACTGCTTCCGGAAGCACCCCGAGGCTACTCABAATGTGGCTCG 1800
 Qy |||||
 2139 GGGCCCTGGTTGACACTAGTGTGCTAGTATGATACCCATACAGGCTTGGGCTACCCC 2198
 Db |||||
 1801 GGGCCTTGGTTGACACTAGTGTGCTAGTATGATACCCATACAGGCTTGGGCTACCCC 1860
 Qy |||||
 2199 TGCATCTCAATTTTCCATCTTTAAGTTAGGATGATGTGGGGGGGCTGGAGCAGG 2258
 Db |||||
 1861 TGCACTGTCACTCTTCCATCTTTAAGTTAGGATGATGTGGGGGGGCTGGAGCAGG 1920

2259 CTCAATGCGCGCATGCAATTGGACTCGAGGAGCGCTGTAACTTGGAGCACAGGATAGG 2318
 Db |||||
 1921 CTTAATGCTGCAATGCAACTGGACCCGAGGAGAGCGTTGCACTTGGACACAGGAGCAGA 1980
 Qy |||||
 2319 TCAGAACTCAGCCCGCTGCTGTCTCTCAACAGAGTGGCAGATACGCTGCTGTGCTTTC 2378
 Db |||||
 1981 TCGGAGCTCAGCCCGCTGCTGTCTCTCAACAGAGTGGCAGGTTCTGCCCTGCTCTTTC 2040
 Qy |||||
 2379 ACACCCCTACCGGCTTATCCACTGGTTGATCCATCTCCATCAGAACATCGTGGACGTG 2438
 Db |||||
 2041 ACCACCCCTACCGGCTTATCCACTGGTTGATCCACCTCCATCAGAACATCGTGGACGTG 2100
 Qy |||||
 2439 CAATACCTCTACCGTGTAGGCTCAGGTTTGTCTCTTTGCAATCAAAATGGGAGTACATC 2498
 Db |||||
 2101 CAATACCTCTACCGTGTAGGCTCAGGTTTGTCTCTTTGCAATCAAAATGGGAGTATGTC 2160
 Qy |||||
 2499 CTGTTGCTTTTCTCTCTGCGAGACGCGCGGTGTGCTGCTGCTGTTGGATGATGCTG 2558
 Db |||||
 2161 GTGTTGCTTTTCTCTCTGCGAGCGCGCTGTGTCCTGCTGTTGGATGATGCTG 2220
 Qy |||||
 2559 CTGATAGCCAGGCTGAGCCGCTTAGAGAACTTGGTGGTCTCAATGGGGCTCGGTG 2618
 Db |||||
 2221 CTGATAGCCAGGCTGAGCCGCTTAGAGAACTTGGTGGTCTCAATGGGGCTCGGTG 2280
 Qy |||||
 2619 GCGGAGCGCATGTTATCTCTCTTTTGTGTTCTCTGCGCCGCTGCTGCTAATTAAG 2678
 Db |||||
 2281 GCGGAGCGCATGTTATCTCTCTTTTGTGTTCTCTGCGCCGCTGCTGCTAATTAAG 2340
 Qy |||||
 2679 GCGGAGCGCTGCTCTGCGGCGGCTATGCTTTTATGGCGTATGGCGCTGCTGCTGCTC 2738
 Db |||||
 2341 GCGGAGCTGCTCTCTGCGGCGGCTATGCTTTTCTATGGCGCATGGCGCTGCTGCTGCTC 2400
 Qy |||||
 2739 CTACTGGGCTTACCAACAGCTTACGCTTGGACGCGGAGATGCTGCTGCTGCTGCTG 2798
 Db |||||
 2401 CTCTTGACATTTACCAACAGCTTACGCTTGGACGCGGAGATGCTGCTGCTGCTGCTG 2460
 Qy |||||
 2799 GGTGCGGCTTCTGTTAGTGTGTTTGTGACCTTGTGACCTTGTACCACTACAAAGTGTTC 2858
 Db |||||
 2461 GGTGCGGCTTCTGTTAGTGTGTTTGTGACCTTGTGACCTTGTGACCTTGTGACCTTGT 2520
 Qy |||||
 2859 ACTAGGCTCATATGTTGTTTACAAATCTTTATCACCAGAGCCGAGGCGCATGCAAGTG 2918
 Db |||||
 2521 GCTAGGCTCTATGTTGTTTACAAATCTTTATCACCAGAGCTGAGGCGCATGCTGCTG 2580
 Qy |||||
 2919 TGGGTCCTCCCTCAAGCTTGGGAGGCGCGATGCGCATCATCTCTCTCAAGTGTGCG 2978
 Db |||||
 2581 TGGGTCCTCCCTCAAGCTTGGGAGGCGCGATGCGCATCATCTCTCTCAAGTGTGCG 2640
 Qy |||||
 2979 GTTATCCAGAGTAAATTTTGTGATCATCAAACTCTGCTGCTGCTGCTGCTGCTGCTG 3038
 Db |||||
 2641 GTCCACCCAGAGTAAATTTTGTGATATCAAACTCTGATGCTGCTGCTGCTGCTGCTG 2700
 Qy |||||
 3039 ATGCTGCTCCAGGCTGCGCATACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3098
 Db |||||
 2701 ATGCTGCTCCAGGCTGCGCATACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
 Qy |||||
 3099 CGTGCATGCTATGTTAGTGGGAAAGTGGCGGGGCTCATTTATGTCGCTGCTGCTGCTG 3158
 Db |||||
 2761 CGTGCATGCTATGTTAGTGGGAAAGTGGCGGGGCTCATTTATGTCGCTGCTGCTGCTG 2820
 Qy |||||
 3159 AAGTGGGCGGCTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3218
 Db |||||
 2821 AGACTGGGCGGCTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
 Qy |||||
 3219 GCGGAGCGGCTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3278
 Db |||||
 2881 GCGGAGCGGCTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940
 Qy |||||
 3279 GAGACCAAGTTCATCACTGGGAGGAGACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3338
 Db |||||
 2941 GAGACCAAGTTCATCACTGGGAGGAGACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
 Qy |||||
 3339 CTACCCGCTCTCCGCCCCGAGGAGATATTTTGGGACCGGCTGCTGCTGCTGCTGCTG 3398

Db 3001 |||||CTACCTGTCTCCGCCGAAGGGAGGAGATACCTCTGGGCGCGCGATAGTCTAGTA 3060
QY 3399 GGGCAAGGGTGGCGACTCTCTGGCGCCATCAGGGCCTACTCCCAACAAACCGGGGGCGTA 3458
Db 3061 GGGCAAGGGTGGCGACTCTCTGGCGCCATCAGGGCCTACTCCCAACAAACCGGGGGCGTA 3120
QY 3459 CTTGGTTCATCATCACTAGCTCACAGCGCGGGAAGAACCAGGTGCGAAGGGGAGGTT 3518
Db 3121 CTTGGTTGCATCATCAGAGTCTCAGAGCGCGGAACAAGACCAGGTTCAGGGGAGGTT 3180
QY 3519 CAAGTGGTTCTTACCGCAACACAATCTTTCTCTGGCGACCTGCATCAACGGGCGTGTGCTGG 3578
Db 3181 CAAGTGGTCTCACCGCAACAACAATCTTTCTCTGGCGACCTGCGTCAACGGCGTATGTTGG 3240
QY 3579 ACTGTCTACATGGCGCTGGCTCGAAGACCCCTAGCCGGTCCAAAGAGTCCAAATCACCCAA 3638
Db 3241 ACTGTCTACATGGTGTGCTTCAAGAGCTCTAGCCGGCCCAAAAGGCGCAATCGCCAG 3300
QY 3639 ATGTACACCAATGTAGACTTGGACCTCGTCCGCTGGCAGGCGCCCGCGGGCGCGCTCC 3698
Db 3301 ATGTACACTAATGTAGACCAAGATCTCGTGGCTGGCGCGCGCCCGCGGGCGCGTTCC 3360
QY 3699 ATGACACCATGCGAGCTGTGGAGCTCGGACCTTTACTTTGGTCACGAGACATGCTGATGC 3758
Db 3361 CTGACACCATGCACTGTGGCAGCTCGGACCTTTACTTTGGTTACGAGACATGCGATGTT 3420
QY 3759 ATTCCGGTGGCGCGGAGCGACAGCGGGAAGTCTACTCTCCCGCAGGCGCGTCTCC 3818
Db 3421 ATTCCGGTGGCGCGGCGGACAAATAGAGGAGCTTGTCTCCCGCAGGCGCTGTCTCC 3480
QY 3819 TACCTCAAGAGCTCTCGGTGGTCCATTGCTTTCCCTTCCGTCGCGGACGCTGTGGGCGTC 3878
Db 3481 TACTTTGAAGGGCTCTTCGGGTGGCCCACTGCTCTGCCCTTCGGGGCAGCGTGTGGGCGTC 3540
QY 3879 TTCCGGGCTGCTGTGTGCAACCGGGGGTGGCAAGGGGGTGGACTTCATACCCGTTGAG 3938
Db 3541 TTCCGGGCGCTGTATGCACCCGGGGGTGGCAAGGGGGTGGATTTGTCCCGTTCAG 3600
QY 3939 TCTATGGAAATPACATGCGGTCTCCGFTCTTCAAGACAACTCAACCCCGCGCTGTA 3998
Db 3601 TCCATGGAAATPACTATGCGGTCCCGGTCTTCAAGACAACTCATCTCCCGCGCGTA 3660
QY 3999 CGCAGACATTCGAAGTGCACTGTGACGCTCTCTAGCAGCACTCAACCCCGCGCTGTA 4058
Db 3661 CGCAACAACTTCAAGTGCGCCATCTACACGCTCCCACTGGCAGCGGCAAGAGCACTAGA 3720
QY 4059 GTGCGGCTGCTATGTCAGCCCAAGGTACAAGGTGCTGCTCTGAAACCCGTCCTGGCC 4118
Db 3721 GTGCGGCGCATATGCGGCCCAAGGTTACAAGTGTGTTCTGAAACCGTCTGTTGCC 3780
QY 4119 GGCACCTTAGGTTTGGGCGTATATGTCCAAAGGCAACGGTATGCAACCTTAACATCAGA 4178
Db 3781 GCTACCTTAGGTTTGGGCGTATATGTCTAAGGACACATGGTACCGACCTTAACATCAG 3840
QY 4179 ACTGGGTAAAGCACTTACACGCGCGCTCCATTACGTACTCCACCTATGGCAAGTTC 4238
Db 3841 ACTGGGGTAAGGACCAATTAACGCGCGCCCACTTACGTACTCCACCTATGGCAAGTTC 3900
QY 4239 CTTGGCGAGCGGTGGTGTCTGGGGGCGCTATGACATCATTAATATGTGATGAGTGGCCAC 4298
Db 3901 CTTGGCGAGCGGTGGTGTCTGGGGGCGCTTACGACATCATTAATGTGCGATGAGTGGCAC 3960
QY 4299 TCAACTGACTCGACTACCATCTTGGGCATCGGCACAGTCTCGAACAGCGGAGACGGCT 4358
Db 3961 TCAACTGACTCAACTACTATCTTGGGCATCGGCACAGTCTCGAACAGCGGAGACGGCT 4020
QY 4359 GGAGCGGCTGTGCTGTCTCGCCACCGCTACACCTCCGGGATCGGTTACCGTCCGACAC 4418
Db 4021 GGAGCGGCTGTGCTGTCTCGCCACCGCTACGCTTCAGGATCGGTACCGTTCGACAC 4080
QY 4419 CCCAATATCGAGGAATAGGCTGTCCAAACAATGGAGAGATCCCTTCTATGGCAAGCC 4478
|||||

Db 4081 CCCAATATCGAGGAGTGGCCTGTGCAACACTGGAGAGATCCCTTCTACGGCAAGCC 4140
QY 4479 ATCCCAATTCAGGCATCAAGGGGGAGGCATCTCAATTTCTGCCAATCCCAAGAAGAA 4538
Db 4141 ATCCCAATCGAAGCCATCAAGGGGGAGGACCTCAATTTCTGTCACTCCAAAGAGAG 4200
QY 4539 TGTGACGAGTCCGCGCAAGCTGACAGGCTCGGACTGAAACGCTGTAGCATATTTACCGG 4598
Db 4201 TGGCAGAGCTTGGCGCAAGCTGTAGGCTCGGAATCAATGCTGTAGCGTATTTACCGG 4260
QY 4599 GGCCTTGATGTGCTGTCATACCGCTTATCGGAGACGTGTTGTGTGGCAACAGACGT 4658
Db 4261 GGTCTTGATGTGCTGTCATACCGACAGGGGAGACGTGTTGTGTGGCAACAGACGT 4320
QY 4659 CTAATGACGGTTTACCGGCGATTTTGACTCAGTGATCGACTGCAATACATGTGTCAAC 4718
Db 4321 CTAATGACGGCTATACCGGTGACTTTTGATTCAGTGATCGACTGAATACGTGTGTCAAC 4380
QY 4719 CAGACAGTTCGACTTTGAGTTGGATCCCACTTCCACATTTGAGACGACGACCGTCCCCAA 4778
Db 4381 CAGACAGTTCGACTTTCAGTTGGACCCCACTTCCACATTTGAGACGACGACCGTCCCCAA 4440
QY 4779 GACCGGTGTCGCGCTCGCAACGGCGAGGTAGAACTGGCAGGGTAGGAGTGGCATCTAC 4838
Db 4441 GACGAGTGTGCGCTCGCAGCGCGGGTAGGACTGGCAGGGGACGAGGGGGGATATAC 4500
QY 4839 AGGTTTGTGACTCCAGGAGAACGCGCTCGGGCATGTTCCGATTTCTCGGTCTGTGTGAG 4898
Db 4501 AGGTTTGTAACTCCGGGGGAAACGCGCTCGGGCATGTTCCGATTTCTCGGTCTGTGTGAG 4560
QY 4899 TGTATGACCGGGCTGTGCTTGGTATGAGTCTACGCGCGCTGAGACCTCGGTTAGGTTG 4958
Db 4561 TGTATGACCGGGCTGTGCTTGGTATGAGTCTACCGCTCACCCCGCTGAGACCTCGGTTAGGTTG 4620
QY 4959 CGGCTTACCTAAATAACACAGGGTTCGCGTCTGCGAGGACCAATCTGGAAGTCTCTGGAG 5018
Db 4621 CGGCTTACCTAAATAACACAGGATTCGCGTTTTCGAGGACCAATCTGGAAGTCTCTGGAG 4680
QY 5019 AGCCTTTCACAGGCTCACCCATAGATGCCACATCTCTGTCGCCAGACTTAACAGGCA 5078
Db 4681 AGCCTTTCACAGGCTCACCCATATAGATGCCACATCTCTGTCGCCAGGAGGCA 4740
QY 5079 GGAGCAACTTTCCTTACCTGGTGGCATATCAAGCTACAGTGTGCGCGCAGGCTCAAGCT 5138
Db 4741 GGAGTAACCTTCCCTTACCTGGTGGCATATCAAGGCAAGTGTGCGCCAGGCTCAGGCC 4800
QY 5139 CCACTCCATCGTGGGACCAAAATGTGGAAGTGTCTATACGGCTGAACCTTACACTGCAC 5198
Db 4801 CCACTCCATCGTGGGATCAAAATGTGGAAGTGTCTATACGGCTTAAACCCACGCTGCAC 4860
QY 5199 GGGCCACACCCCTGCTGTATAGGTAGGAGCGCTCCAAAATAGGTCACTCTCACACAC 5258
Db 4861 GGGCCACACCCCTGCTGTATAGGTAGGAGCGCTCCAAAATAGGTCACTCTCACACAC 4920
QY 5259 CCCATAACTAAATACATCATGGCATGTGCGGTGACCTCGGAGGTGCTCACTAGCACCC 5318
Db 4921 CCCATAACCAATACATCATGGCATGTGCGCGGACCTGGAAGTCTCACCAGCACCC 4980
QY 5319 TGGGTGCTGTAGGCGGAGTCTTGGAGCTTGGCGCGCATACTGCTGACAGACGAGT 5378
Db 4981 TGGGTGCTGTAGGCGGAGTCTTGGAGCTTGGCGCGCATATTTGCTTGACAAACAGGCACT 5040
QY 5379 GTGTCATTTGGGCGAGGATCATCTTGTCCGGGAGGACGAGTGTGTTCCCGGACAGGAA 5438
Db 5041 GTGTCATTTGGGCGAGGATCATCTTGTCCGGGAGGCGGCTGTGTTCCCGGATAGGAA 5100
QY 5439 GTCTCTTACCAGGAGTTCGATGAGATGGAAGTGTGCTTCACTCACTTCTTCTTACATCGAG 5498
Db 5101 GTCTCTTACCAGGAGTTCGATGGAATGGAAGTGTGCGCTCGCACCTCTCTTACATCGAA 5160
QY 5499 CAGGAATGAGTGTGCGCGAGCAATTCAGCAAAAGGCGCTCGGGTGTGTCGAAACGGCC 5558
Db 5161 CAGGAATGCAACTCGCGGAGCAATTTCAAGCAGAGGCGCTCGGGTGTGTCGAAACAGCC 5220
|||||

Qy	5559	ACCAAGCAACGAGGCTGCTCCGCTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACC	5618
Dd	5221	ACCAAGCAGCGAGGCTGCGCTCCGCTGGTGGAGTCCAAGTGGCGAGCTTTGGAGACC	5280
Qy	5619	TTCTGGCGGAAGCACATGTGGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCC	5678
Dd	5281	TTCTGGCGAAAGCACAAAGTGGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCC	5340
Qy	5679	ACTCTGCTGGAAACCCCGATAGCATCAATGATGGCAATTTACAGCTTCTATCACTAGC	5738
Dd	5341	ACCCTGCTGGGAACCCCGGATAGCATCACTGATGGCATTCACAGCTTCTATCAACAGC	5400
Qy	5739	CGCTCAACCAACCAACACCCCTCTGTTTAAATCTTTGGGGGATGGGTGGCTTGCCCAA	5798
Dd	5401	CGCTCAACCAACCAACACCCCTCTGTTTAAATCTTTGGGGGATGGGTGGCTTGCCCAA	5460
Qy	5799	CTGCTCTCTCCAGCGCTGGTCAGCTTTCGTGGGCGCGGCATCGCGAGCGGCTGTT	5858
Dd	5461	CTGCTCTCCCGCAGCGCTGCTCGGCTTTCTGTTGGGCGCTGGTATCGCTGGTGGCTGTT	5520
Qy	5859	GGCAGCATAGGCTTTGGAAAGTGTCTCTGGACATCTTGGCGGCTATGGCGGAGGGGTA	5918
Dd	5521	GGCAGCATAGGCTTTGGAAAGTGTCTAGTGGACATCTTGGCGGCTATGGCGGAGGGTG	5580
Qy	5919	GCCGGGCACTCGTGGCTTTAAAGTTCATGAGCGGCGAGTGGCTTCCACCGAGGACCTG	5978
Dd	5581	GCTGGCGGCTCGTGGCTTTCAAGGTCATGAGCGGCGAGGCGCTCTGCGCGAGGACCTG	5640
Qy	5979	GTCAACTTACTCCCTGCGCATCTCTCTCTGTTGGTGGCTGCTGCGGGTCTGTGGCA	6038
Dd	5641	ATCAATTTGCTGCTGCGCATCTCTCTCTGTTGGTGGCTGCTGCGGGTCTGTGTGCA	5700
Qy	6039	GCAATCTGCTGCGCACGTGGGCGCGGAGAGGGGCTGTGCACTGGATGGAACCGGCTG	6098
Dd	5701	GCAATCTGCTGCGCATGTGGGCGCGGAGAGGGGCGCTGCACTGGATGGAACCGGCTG	5760
Qy	6099	ATAGCGTTTCGCTTCCGGGGTAAACACGTCTCCCTACGCACTATGTGCCCTGAGAGCGAC	6158
Dd	5761	ATAGCGTTTCGCTTCCGGGGTAAACCATGTCTCCCGCCAGCACTATGTGCCCTGAGAGCGAC	5820
Qy	6159	GCTGACGACGTGTCACTCAGATCCTCTTAGCTTACCATCACTCAACTGCTGTAAGCGG	6218
Dd	5821	GCGCAGCGGCTGTCACTCAGGCTCTCTCCAGCTTACCATCACTCAACTGCTGTAAGCGG	5880
Qy	6219	CTCCACAGTGGATTAATGAGGACTGTCTACGCGCATGCTCCGCTCGTGGCTTAAGGAT	6278
Dd	5881	CTCCACAGTGGATTAATGAGGACTGTCTACGCGCGTCTCCGCTCGTGGCTGAGGAT	5940
Qy	6279	GTTTGGGATGGATATGACCGGTTGTGACTGTCAAGACCTGGCTCCAGTCCAAACTC	6338
Dd	5941	GTTTGGGACTGGGTGTGACCGGTTGTGAGTCACTTCAAGACCTGGCTCCAGTCCAAAGTC	6000
Qy	6339	CTGCGCGGTTACCGGAGTCCCTTTCTGTCTATGTCGCAACCGCGGTACAAGGAGTCTGG	6398
Dd	6001	CTGCGCGGTTACCGGAGTCCCTTTCTGTCTATGTCGCAACCGGTACAAGGAGTCTGG	6060
Qy	6399	CGGGGAGCGCATATGCAACCACTGCGCCATGCGGAGCACAGATCGCGGACATGTC	6458
Dd	6061	CGGGGAGCGCATATGCAACCACTGCGCCATGCTGAGACAGATCGCGGACATGTC	6120
Qy	6459	AAAAAGGTTCCATGAGGATCGTAGGGCTTAGAACCTTCAGCAACACGTTGGCAACGGAACG	6518
Dd	6121	AAAAAGGTTCCATGAGGATCATCGGCGGGAACCTTCAGCAACACGTTGGCAACGGAACG	6180
Qy	6519	TTCCCCATCAACGATACACCAAGGACCTTTGCAACCCCTCCCGGCGCCCAATATTCC	6578
Dd	6181	TTCCCCATCAACGATACACCAAGGACCTTTGCAACCCCTCCCGGCGCCCAATATTCC	6240
Qy	6579	AGGGCGCTATGGCGGCTGCTGAGGAGTACGTGGAGGTTACCGCTGTGGGGATTTTC	6638
Dd	6241	AAGGCGCTGTGGCGGCTGCTGAGGAGTACGTGGAGGTTACCGCTGTGGGGATTTTC	6300

Qy	6639	CACTAGTGACGGGATGACCACTGACAAAGTAAAGTGCCTGCGAGGTTCCGGCCCCC	6698
Dd	6301	CACTAGTGACGGGATGACCACTGACAAAGTAAAGTGCCTGCGAGGTTCCAGTCTCT	6360
Qy	6699	GAATTTTTCACGAGGATGAGTGGGTGTCACAGGTACGCTCCGGCGTCAAAACCT	6758
Dd	6361	GAGTTTTTTCACGAGGATGAGTGGGTGTCACAGGTACGCTCCGGCGTCAAAACCT	6420
Qy	6759	CTTTACGGGAGGACGTCAGTTCCAGGTGGGCTCAACCAATACTTGGTGGGTGCGAG	6818
Dd	6421	CTTTACGGGAGGATGAGTTGTTTCCAGGTGGGCTCAATCAATACCTGGTGGGTGCGAG	6480
Qy	6819	CTCCCATGCGAGCCCAACCGGACGTAAACAGTCTTACTTCCATGTCTACCGATCCCTCC	6878
Dd	6481	CTCCCATGCGAGCCCAACCGGACGTAAACAGTCTTACTTCCATGTCTACCGATCCCTCC	6540
Qy	6879	CACATTACAGCAGAGACGGCTAAGCGTAGGCTGGCTAGAGGGTCTCCCGCTCTTTAGCC	6938
Dd	6541	CACATTACAGCAGAGGCGCTAAGCGTAGGCTGGCTAGAGGGTCTCCCGCTCTTTAGCC	6600
Qy	6939	AGCTCATAGCTAGCCAGTTGCTGGGCTTCTTGAAGCGGACATGCATACCC-----	6993
Dd	6601	AGCTCTTACGTAGCCAGCTGCTGGGCTTCTTGAAGCGGACATGCATACCCATTCT	6660
Qy	6994	-----ACCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCTTGTGGCGGAG	7046
Dd	6661	TCCTATAATCTTGACTCTCCGACGTGACCTCATTTGGCGGCCAACCTTGTGGCGGAG	6720
Qy	7047	GAGATGGCGGAAACATCACTCGCTGGAGTCAGAGAAATAGGTAGTAAATTTGGACTCT	7106
Dd	6721	GAGATGGCGGAAACATCACTCGCTGGAGTCAGAGAAATAGGTAGTAAATTTGGACTCT	6780
Qy	7107	TTCGAACCGCTTACGCGGAGGGGATGAGAGGAGATATCCGTCGCGGCGGAGATCTTG	7166
Dd	6781	TTCGAGCGCTTACGCGGAGGGGATGAGAAATATCCATTTGGCGGCGAGATCTTG	6840
Qy	7167	CGAAATCCAGGAAGTTCCCTCAGCGTTGCCATATGGGACGCGCCGGACTCAATCTCT	7226
Dd	6841	CGAAAGTCCAAGAGTTCCCGCGGCGATACCATATGGGACGCGCCGGATTAACATCTCT	6900
Qy	7227	CCACTGTAGAGTCTTGAAAGGACCCCGACTAGCTCCCTCCGGTGGTACAGGATGCCCA	7286
Dd	6901	CCATTGTAGAGTCTTGAAAGACCCCGACTAGCTCCCTCCGGTGGTACAGGATGCCCA	6960
Qy	7287	TTGCGACCTACCAAGGCTCTCCAATACCACTTCCACGAGAAAGAGGACGGTTGTCTG	7346
Dd	6961	TTGCGACCTCTCAAGGCGCTCCAATACCACTTCCACGAGAAAGAGGACGGTTGTCTG	7020
Qy	7347	ACAGAAATCCAAATGTGTTCTGCTTGGCGGAGCTCGCCATAGAGCTTTCGGTAGTCC	7406
Dd	7021	ACGAGCTCCACCGTGTCTTCTTTTGGCGGAGCTCGCTACCAAAACCTTCGGGAGCTCC	7080
Qy	7407	GGATCGTGGCGGTTGATAGCGGACGCGGACCGCTTCTGACCTGGCTCCGACGAC	7466
Dd	7081	GAATTTGTCGCGCGGACGAGGACGCGGACCGGCTCTCTGACGAGACCTTCGACAAAC	7140
Qy	7467	GGTGACAAAGGATCCGAGCTTGTAGTCTGTTCTTCCATGCCCCCTTTGAAGGGAGCGG	7526
Dd	7141	GGCGGCAAGACTCCGAGCTGAGTCATGCTCTCTATGCCCCCTTTGAGGGGAGCGG	7200
Qy	7527	GGGAGCCCGATCTCAGCGAGCGGTCTTGTCTACCGTGTGAGGAGGCTAGTGGAGAT	7586
Dd	7201	GGGAGCCCGATCTCAGCGAGCGGTCTTGTCTACCGTGTGAGGAGGCTAGTGGAGAT	7260
Qy	7587	GTGCTGTGCTCAATGTCTTATACGTGGACGCGCCCTGATCACGCCATGGCTGCG	7646
Dd	7261	GTGCTGTGCTCAATGTCTTATACGTGGACGCGCCCTGATCACGCCATGGCTGCGCGG	7320
Qy	7647	GAGGAAAGTAAAGTGTGCGCATCAACCGGTTGAGCAACTCTTTGCTGCTGCTACCAACATG	7706
Dd	7321	GAAGAAAGCAAGTGTGCGCATCAACCGGTTGAGCAACTCTTTGCTGCTGCTACCAACATG	7380
Qy	7707	GTCTACGCCCAACATTCGCGGAGCGCAAGCTTCGCGAGAAAGAGGTACCTTTGACAGA	7766

; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 9275
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding hepatitis C virus polypeptide
; OTHER INFORMATION: derived from HCV-N
US-11-006-313-39

Query Match 82.2%; Score 7885.8; DB 10; Length 9275;

Best Local Similarity 91.0%; Pred. No. 0;

Matches 8437; Conservative 0; Mismatches 817; Indels 21; Gaps 4;

QY	342	ATGAGCAGCAATCTTAAACCTCAAGAAACCAACAAAGCTAAACCAACCGCCGCCACAG	401
DB	1	ATGAGCAGCAATCTTAAACCTCAAGAAACCAACAAAGCTAAACCAACCGCCGCCACAG	60
QY	402	GACGTCAAGTTCGGGGCGGTGTCAGATCGTTGGTGGAGTTTACCTGTTGCCGCGCAGG	461
DB	61	GACGTCAAGTTCGGGGCGGTGTCAGATCGTTGGTGGAGTTTACCTGTTGCCGCGCAGG	120
QY	462	GGCCCGAGGTGGGTGTCGGCGGACTAGGAAGGCTTCCGAGCGGTCCGAACCTCGTGGGA	521
DB	121	GGCCCGAGGTGGGTGTCGGCGGATCAGGAGACTTCCGAGCGGTCCGAACCTCGTGGGA	180
QY	522	AGCGCAGCAACTTATCCCAAGGCTCGCGACCCGAGAGGCGAGGCTGGGTCTCAGCCCGGG	581
DB	181	AGCGCAGCAACTTATCCCAAGGCTCGCGGCGCGGAGGCGAGGCTGGGTCTCAGCCCGGG	240
QY	582	TACCTTGGCCCTCTATGCAATGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT	641
DB	241	TATCTTGGCCCTCTATGCAATGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT	300
QY	642	CGCGGCTCCCGGCTAGTGGGGCCCAAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	701
DB	301	CGCGGCTCCCGGCTAGTGGGGCCCAAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	360
QY	702	AAGGTATCATGATACCTTATATGCGGCTTCGCGAGCTCATGAGGCTCATTCGCTCGTC	761
DB	361	AAGGTATCATGATACCTTATATGCGGCTTCGCGAGCTCATTCGCTCGTC	420
QY	762	GGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGCGACACGCTGTCGGGTTCTGGAGGAC	821
DB	421	GGCGCCCGCTAGGGGGCGCTGCCAGGCGCTTGCGACACGCTGTCGGGTTCTGGAGGAC	480
QY	822	GGCGTGAATATGCAACAGGGAACTTGCCCGGTTGCTCTTCTCTATCTTCTTGGCT	881
DB	481	GGCGTGAATATGCAACAGGGAACTTGCCCGGTTGCTCTTCTCTATCTTCTTGGCT	540
QY	882	CTGCTGTCTCTTGAACCATCCAGCTTCCGCTTATGAAGTGCACACGCTGTCGGGATA	941
DB	541	CTGCTGTCTCTTGAACCATCCAGCTTCCGCTTATGAAGTGCACACGCTGTCGGGATA	600
QY	942	TACATGTCAAGAACGATGTCCTCAACTCAAGCAATGTTGATGAGGACGCGGACGTCATC	1001
DB	601	TACATGTCAAGAACGATGTCCTCAACTCAAGCAATGTTGATGAGGACGCGGACGTCATC	660
QY	1002	ATGCATATCTCCGGGTGGTGGCTGTTTTCAGGAGGTTACAGCTCCCGTGGTGGTA	1061
DB	661	ATGCATATCTCCGGGTGGTGGCTGTTTTCAGGAGGTTACAGCTCCCGTGGTGGTA	720
QY	1062	CGGCTCACTCCAGCTCGCGGCGAGGAATGCGACGCTCCCACTACGACAAATACGACGC	1121
DB	721	CGGCTCACTCCAGCTCGCGGCGAGGAATGCGACGCTCCCACTACGACAAATACGACAC	780
QY	1122	CACGTGCACTTGTCTGTTGGGACGGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTC	1181
DB	781	CACGTGCACTTGTCTGTTGGGACGGCTGCTTCTGCTCCGCTATGTACGTGGGGATCTC	840

QY	1182	TGCGGATCTATTTTCTCTCTCCAGCTGTTTACCTTCTCGCTCGCGGCATGAGACA	1241
DB	841	TGCGGATCTATTTTCTCTCTCCAGCTGTTTACCTTCTCGCTCGCGGCATGAGACA	900
QY	1242	GTGAGGACTGCAACTGCTCAATCTATCCGGGCATGATATCAGGTCAACCGCATGGCTGG	1301
DB	901	TTGAGGACTGCAATTTGTTGATCTATCCCGGCACGCGTCAGGTCAACCGCATGGCTGG	960
QY	1302	GATATGATGATGAATCTGTCACCTCAACAGCCCTAGTGTGTCGAGTGTCTCGGATC	1361
DB	961	GACATGATGATGAATCTGTCACCTCAACAGCCCTGTCAGTGTCTCGGATGTCTCGGATC	1020
QY	1362	CCACAGCTGTGTCGACATGTCGGGGGCGCCACTGCGGAGTCTCTGGGGGCGCTTGGC	1421
DB	1021	CCACAGCTGTGTCGACATGTCGGGGGCGCCACTGCGGAGTCTCTGGGGGCGCTTGGC	1080
QY	1422	TACTATTTCATGTCAGGAACTGGGCTAAGGTTCTGATTTGGGCTTACTCTTTTGGCGGC	1481
DB	1081	TACTATTTCATGTCAGGAACTGGGCTAAGGTTCTGATTTGGGCTTACTCTTTTGGCGGC	1140
QY	1482	GTTGACGGGAGACCCACACGACGGGGAGGTGGCGGCGCACACCTCCGGGTTCAAG	1541
DB	1141	GTTGACGGGACACCCCTCAACAGGGGGGACGCTGCCCGCTCACAGCGGTTTCCGC	1200
QY	1542	TCCCTTTTCTATCTGGGGCTCTCAGAAATCCAGTTTGTGATATCAACAGCGCAGCTGG	1601
DB	1201	GGCCTCTTTTACCTGGGGCTCTCAGAAATCCAGTTTGTGATATCAACAGCGCAGTTGG	1260
QY	1602	CACATCAACAGGACTGCCCTAAATTCGAATGACTTCCCTCCAACTGGGTTCTTTGGCGCG	1661
DB	1261	CACATCAACAGGACTGCCCTGAAATTCGAATGACTTCCCTCCAACTGGGTTCTTTGGCGCG	1320
QY	1662	CTGTTTACACACACAAAGTTCAACTCTGTCGGGTGCCCGGAGCGCATGGCAGCTGGCG	1721
DB	1321	CTGTTTACACACACAAAGTTCAACTCTGTCGGGTGCCCGGAGCGCATGGCAGCTGGCG	1380
QY	1722	CCATTTGACTGTGTCAGGCTGAGGCTGGGCGCCCATCACTTATCTAAGCCT---AACAGC	1778
DB	1381	TCCATTTGACAAAGTTTCGACAGGAGTGGGCTCTATCACTTATGCTGAGCCTTACAAAGAC	1440
QY	1779	TCCGATCAGAGGCTTATGCTGSCATTACGCGCTCGACGCTGCTGCTGCTACCCGCG	1838
DB	1441	CCGACACAGAGGCTTATGCTGSCATTACCCACTTCAACAAATGCTGATCTGATCTGGC	1500
QY	1839	TCGAGGCTGTGTCAGTGTATTTGTTTACCCCAAGCCCTGTTGTTGGTGGGACCAACC	1898
DB	1501	TCGAGGCTGTGTCAGTGTATTTGCTTCAACCCCAAGCTCTGTTGCTGTTGGGACCAACC	1560
QY	1899	GATCGTTCCGGTGTCCCTACGTATAGCTGGGGGAGAAATGAGACAGAGCTGATCTCTC	1958
DB	1561	GATCGTTCCGGCAACCCCTACGTATAGCTGGGGGAGAAACGATCTGAGCTGCTGCTCTT	1620
QY	1959	AACAAACGCGTCCGACACAGGCACTGGTTCCGCTGTATCATGATGAATAGTACTGGG	2018
DB	1621	AACAAACGCGGCGCGCAAGGCACTGGTTCCGCTGTATCATGATGAATAGTACTGGG	1680
QY	2019	TTCACTAAGACGTCGCGAGGTCCCGCTGTAAATCGGGGGGCTCGGTAAACCGCACCTTG	2078
DB	1681	TTCACTAAGACGTCGCGGGGCGCCCGCTGTAAATCGGGGGGCTCGGTAAACCACTTG	1740
QY	2079	ATCTGCCCGCAGGACTGCTTCCGGAAGCACCCCGAGGCTACTTACAAAAATGTTGGTCTG	2138
DB	1741	ACCTGCCCGCAGGACTGCTTCCGGAAGCACCCCGAGGCTACTTACAAAAATGTTGGTCTG	1800
QY	2139	GGGCGCTGTTTGAACCTAGTGTCTAGTACTACCATACAGGCTTTGGCACTACCCC	2198
DB	1801	GGGCGCTGTTTGAACCTAGTGTCTAGTACTACCATACAGGCTTTGGCACTACCCC	1860
QY	2199	TGCACTCTCAATTTTCCATCTTTAAGGTTAGGATGTATGTGGGGGCGTGGAGCACAGG	2258
DB	1861	TGCACTGTCAATTTCTCATCTTTAAGGTTAGGATGTATGTGGGGGCGTGGAGCACAGG	1920

Qy	2259	CTCAATGCCCGCATGCAATTTGGATCTCAGAGAGAGCGGTGTAACTTTGGAGACACAGGATAGG	2318
Db	1921	CTTTAAATCTGCATGCAACTGGACCCGAGGAGAGCGTTTGAACCTTTGGACGACAGGACAGA	1980
Qy	2319	TCAGAACTCAGCCGCTGCTGTCTTACAAACAGAGTGGCAGATACTGCCCTGTGCTTTC	2378
Db	1981	TCGGAGCTCAGCCGCTGCTGTCTTACAAACAGAGTGGCAGGTCTTGCCTGCTCTTTTC	2040
Qy	2379	ACCAACCTACCGGCTTTATCCACTGGTGTTCATCCATCTCCATCAGAAACATCTGTGGACGCTG	2438
Db	2041	ACCACCTACCGGCTGTCTTCCACTGGTGTTCATCCATCAGAAACATCTGTGGACGCTG	2100
Qy	2439	CAATACCTGTACCGGTGTAGGGTGTGCTCTCTCTTTCGAATCAAAATGGGAGTACATC	2498
Db	2101	CAATACCTGTACCGGTGTAGGGTGTGCTCTCTCTTTCGAATCAAAATGGGAGTATGTC	2160
Qy	2499	CTGTTGCTTTTCTTCTTCTTGGCAGACGCGCGGTGTGCTGCTGTGGATGATGCTG	2558
Db	2161	GTGTTGCTTTTCTTCTTCTTGGCGGACGCGCGGTGTGCTGCTGTGGATGATGCTG	2220
Qy	2559	CTGATACCCAGGCTGAGGCGCCTTAGAGAACTTTGGTGTCTCTCAATGGCGGTCCGCTG	2618
Db	2221	CTGATACCCAGGCGGAGGCGCCTTAGAGAACTTGGTGTCTCTCAATGGCGGTCCGCTG	2280
Qy	2619	GCCGGAGCGCATGGTATTTCTCTCCTTTCTTTGTGTCTTCTTGCSCGCGCTGGTACATTAAG	2678
Db	2281	GCCGGAGCGCAGGSCATCCTCTCCTTCTCTGTTCTTCTGTGCGCTTGGTATACATCAAG	2340
Qy	2679	GGCAGGCTGGCTCTCTGGGGCGGCGTATGCTTTTATTTGGCGTGTGCGCTGCTCTCTGCTC	2738
Db	2341	GGCAGGCTGGTCTCTCTGGGGCGGCGTATGCTTTTCTTATGGCGCATGGCGCTGCTCTGCTC	2400
Qy	2739	CTACTGGCGTTTACCACACAGAGCTTACGCTTGACCGGAGAGTGGCTGATCGTGGCGG	2798
Db	2401	CTCTTGACATTTACCACACAGAGCTTACGCGATGACCGGAGAGTGGCTGATCGTGGCGGA	2460
Qy	2799	GGTGGCGTTCTTTGTAGTCTGTGTTATTTGTACCTTTGTCAACATACTACAAAGTGTTTCTC	2858
Db	2461	GGCGCGGTTTTTGTGGGCTGTGGCATTTATGACCTTTGTGCGCATATTACAAAGTGTTCCTC	2520
Qy	2859	ACTAGGCTCATATGGTGGTTTACAATACTTTATCACAGAGCGAGGGCGCATGCAAGTG	2918
Db	2521	GCTAGGCTCTCATGGTGGTTTACAATACTTTATCACAGAGCTGAGGCGCATTTGCAATG	2580
Qy	2919	TGGGTCCCCCCTCAACGTTTCGGGAGGCGCGATGCCATCATCTCTCACTGCTGGCG	2978
Db	2581	TGGGTCCCCCCTCAACGTTCCGGGAGGCGCGATGCCATCATCTCTCACTGCTGGCGA	2640
Qy	2979	GTTTCATCAGAGTTAAATTTTTTGACATCACAAATCTCTGCTCGGCATCTCGGCGCGCTC	3038
Db	2641	GTCCACCAGAGCTAAATCTTTTGATATCACCAAACTTCTGATTGCCATCTCGGACGCTC	2700
Qy	3039	ATGTTGCTCCAGGCTGGCATTAACAGAGTGGCGTACTTGTGCGCGCTCAAGGGCTCATTT	3098
Db	2701	ATGTTGCTCCAGGCTGGCATTAACAGGTTGGCGTACTTGTGTAAGGCGCTCAAGGGCTCATTT	2760
Qy	3099	CGTGCATGCATGTTAGTTCGGAAGTTCGCGGGGTCTATTATGTCCAAATGCTTCTCATG	3158
Db	2761	CGTGCATGCATGTTAGTTCGGAAGTTCGCGGGGTCTATTATGTCCAAATGCGCTTCTCATG	2820
Qy	3159	AAGCTGGCGCGCTGACAGGTACCTGATTTATTAACCATTTACCCCATCTCGCGGACTGG	3218
Db	2821	AGACTGGCGCGCTGACGGGACAGTACGTCTATTAATCACTTCACTCCACCTCGCGGATGG	2880
Qy	3219	GCCACGCGGCGCTACGAGACCTTTCCGTTGGCGGTAGAGCCGCTGCTTCTTCGCGCATG	3278
Db	2881	GCCACGCGGCGCTACGGGACCTTTCCGTTAGCAGTGGAGCCTGCTGCTTCTCTGACATG	2940
Qy	3279	GAGACCAAGTTCATCACTCGGGAGCAGACACCGCTGGTGTGGGACATCATCTTGGT	3338
Db	2941	GAGACCAAGATTCATCACTCGGGGGCGGACACCGCGCGGTGTGGGACATCATCTTGGG	3000
Qy	3339	CTACCGCTCTCCGCGCGGAAGGGGAAGGAGATATTTTTGGGACCGGCTGTAGTCTCGAA	3398

[illegible]

4081	CCCAATATCGAGGAGGTGGCCCTGTGCGAACACTGAGAGATCCCTCTTCTACGGCAAGGCC	4140	DB
4479	ATCCCCATTGAGCCCATCAAGGGGGGAGGCATCTCATTTTCTGCGATTTCGAAGAAGAAA	4538	QY
4141	ATCCCCATCGAAGCCATCAAAGGGGGAGGCACCTCATTTTCTGTCACTCCAAGAAGAAG	4200	DB
4539	TGTGACGAGCTCGCCGCAAAAGCTGACAGGCCTCGGACTGAAAGCTGTAGCATATTTACGG	4598	QY
4201	TGCGACGAGCTTGGCCGAAGCTGTCAAGCCCTCGGAATCAATGCTGTAGCGTATTTACGG	4260	DB
4599	GSCCTTGATGCTGTCCGTGATACCGGCTATCGGAGAGCTGTTGTGCTGGCAACAGACGCT	4658	QY
4261	GGCTTTGATGTGTCCGTATACCGACACCGGAGAGCTGTTGTGCTGGCAACAGACGCT	4320	DB
4659	CTAATGACGGGTTTACC CGCGCATTTTGACTCAGTGATCGACTGCAATACTGTGTACCC	4718	QY
4321	CTAATGACGGGCTATACCGGTGACTTTTGATTCACTGATCGACTGTAATACGTGTGTCA	4380	DB
4719	CAGACAGTCGACTTCAGCTTTGATCCACCTTCAAGATTGAGACGACGACCGTGCCTCCAA	4777	QY
4381	CAGACAGTCGACTTCAGCTTTGAGACCCACCTTCACCATTGAGACGACGACCGTGCCTCCAA	4440	DB
4779	GACCGGTGTCCGGCTCGCAAAGCGGAGCTAGAACTGGCAGAGGCTAGGAGTGCATCTATC	4838	QY
4441	GACGAGTGTCCGGCTCGCAGCGCGGGGTAGGACTGGCAGGGGACAGGGGGGCGATATAC	4500	DB
4839	AGGTTTGTGACTCCAGGAGAAAGCGCCCTCGGGCATGTTCCGATTCTTCGGTCCCTGTGTGAG	4898	QY
4501	AGGTTTGTAACTTCGGGGGAAAGCGCCCTCGGGCATGTTCCGATTCTTCGGTCCCTGTGTGAG	4560	DB
4899	TGCTATGACCGGGCTGTGCTTTGGTATGAGCTCAAGCCCGCTTGAGACCTCGGTAGGTTG	4958	QY
4561	TGCTATGACCGGGCTGTGCTTTGGTATGAGCTCAAGCCCGCTTGAGACCTCGGTAGGTTG	4620	DB
4959	CGGGCTTACTTAAATACACAGGTTGCCCGCTGCGCAGGACCATCTGGAGTCTCTGGGAG	5018	QY
4621	CGGGCTTACTTAAATACACAGGATTGCCCGTTTGCAGGACCATCTGGAGTCTCTGGGAG	4680	DB
5019	AGCGTCTTCAAGGGCTCACCCACATAGATGCCACTTCTGTGCCAGACTAAACAGGCA	5078	QY
4681	AGCGTCTTCAAGGGCTCACCCATATAGATGCCACTTCTGTGCCAGACTAAACAGGCA	4740	DB
5079	GGAGACAACCTTCTTACCTGTGGCATATCAAGCTTACAGTGTGCGCCAGGGCTCAAGCT	5138	QY
4741	GGAGATAAATCTTCCCTTACCTGTGGCATACCAAGCCACAGTGTGCGCCAGGGCTCAGGCC	4800	DB
5139	CGACCTTCATCTGTGGAGCAAAATGTGGAGTGTCTCATACGGCTGAAACCTACACTGCAC	5198	QY
4801	CCACCTTCATCTGTGGGATCAAAATGTGGAGTGTCTCATACGGCTTAAACCCACGCTGCAC	4860	DB
5199	GGGCCCAACCCCTGTGTATAGCTAGGAGCGGTCCAAATCAGGTTCATCTTCAACACAC	5258	QY
4861	GGGCCCAACCCCTGTGTATAGCTAGGAGCGGTCCAAATCAGGTTCATCTTCAACACAC	4920	DB
5259	CCCATAACTAAATACATCATGGCATGTCCGCTGACCTGGAGTTCGTCACTAGCACCC	5318	QY
4921	CCCATAACTAAATACATCATGGCATGTCCGCTGACCTGGAGTTCGTCACTAGCACCC	4980	DB
5319	TGGGTGCTGTAGGCGAGTCTTTCAGCTTTTGGCGGCATCTGCTGTAGCAGCAGGCGAGT	5378	QY
4981	TGGGTGCTGTAGGCGAGTCTTTCAGCTTTTGGCGGCATCTGCTGTAGCAGCAGGCGAGT	5040	DB
5379	GTGGTCAATTTGGGACAGGATCATCTTCTCGGGAGAGCCAGCTGTGCTTCCGACAGGGAA	5438	QY
5041	GTGGTCAATTTGGGAGAGATCATCTTCTCGGGAGAGCCAGCTGTGCTTCCGATAGGGAA	5100	DB
5439	GTCTCTTACCAGGAGTTCGATGAGATGGAAAGTGTGCCTCACAACTTCCTTTACATCGAG	5498	QY
5101	GTCTCTTACCAGGAGTTCGATGAGATGGAAAGTGTGCCTCGACCTCCCTTTACATCGAA	5160	DB
5499	CAGGGAAATGACGTCCCGGACGAATTCAGCAAAAGCGCTTCGGGTGTTGCAAAAGGCC	5558	QY
5161	CAGGGAAATGCACTTCGCGGAGCAATTCAGCAGAAAGCGCTTCGGGTGTTGCAAAAGGCC	5220	DB

Qy	5559	ACCAAGAAAGCGAGGCTGCTGCTCCGCTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACC	5611
Db	5221	ACCAAGCAGCGGAGGCTCCCGCTCCGCTGGTGGAGTCCAAGTGGCGAGCTTTGGAGACC	5280
Qy	5619	TTCTGGCGGAAGCACATGTGGAAATTTTCATCAGCGGAATACAGTACCTACAGGCTTATCC	5678
Db	5281	TTCTGGGCAAGCACAAAGTGGAAATTTTCATCAGCGGATACAGTACTTAGCGGGCTATCC	5340
Qy	5679	ACTCTGCTGGAAACCCCGCATAGCATCATTTAGTGGCAATTTACAGCTTCTATCATAGC	5738
Db	5341	ACCCTGCTGGGAACCCCGCATAGCATCACTGATGGCATTCACAGSCTTATCACCAAGC	5400
Qy	5739	CGGCTACACACCAAAACACCTCTGTTTAAATCATCTTGGGGGAGTGGTGGCTGCCAA	5798
Db	5401	CCGCTACCAACCCAGAAACCCCTCTGTTTAAATCATCTTGGGGGGGTGGGTAGCCGCCAA	5460
Qy	5799	CTGCTCTCCACAGCGCTGCTCAGCTTTCGTGGGCGCGGATCCGCGAGCGGCTGTT	5858
Db	5461	CTGCTCCCCCAGCGCTGCTTCGGCTTTTCGTGGGCGCTGTTATCGCTGGTGGGCTGTT	5520
Qy	5859	GGCAGCATAGGCCCTTGGGAAGTGTCTCGTGGACATCTTGGCGGGCTATGGGCGAGGGTA	5918
Db	5521	GGCAGCATAGGCTTGGGAAGTGTCTAGTGGACATTTCTGGCGGCTATGGGCGAGGGTG	5580
Qy	5919	GCGGGCGCATCTGCTGGCTTTAAAGTTCATAGCGGCGAGGTGCCCTTCACCGAGGACCTG	5978
Db	5581	GCTGGCGGCTCGTGGCGTTTCAAGGTCATGAGCGGGGAGGCGCCCTCTGCCGAGGACCTG	5640
Qy	5979	GTCAACTTACTCCCTGCATCTCTCTCTGCTGTCCTCTGCTGGGGTCTGTCGCGCA	6038
Db	5641	ATCAATTTGCTCCCTGCGCATCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCA	5700
Qy	6039	GCAATACTGCGTGGGCACGTGGGCGCGGGAGAGGGGGCTGTGCAGTGGATGAACCGGCTG	6098
Db	5701	GCAATACTGCGTGGGCATGTGGGCGCGGAGAGGGGGCGTGCAGTGGATGAACCGGCTG	5760
Qy	6099	ATAGCGTTCGCTTCGGGGGTAAACCAAGTCTCCCTTACGCACTATGTGCTCAGAGGCAAC	6158
Db	5761	ATAGCGTTCGCTTCGGGGGTAAACCAAGTCTCCCTTACGCACTATGTGCTCAGAGGCAAC	5820
Qy	6159	GCTGCAGCATGTCACTCAGATCTCTCTAGCCTTACCATCACTCAACTGCTGTAAGCGG	6218
Db	5821	GCCGAGCGGCTGTCACTCAGGCTCTCTCAGGCTTACCATTACCCAGCTGCTTGAAGAGG	5880
Qy	6219	CTCCACAGTGGATTAATAGGACTGCTCTACGCGCATGCTCCGCTCGTGGCTTAAGGGAT	6278
Db	5881	CTCCACAGTGGATTAATAGGACTGTTTACGCGCGTGTTCGCGCTCGTGGCTGAGGGAT	5940
Qy	6279	GTTTGGGATTGGATATGACCGGTGTGACTGTCACTTCAAGACTGGCTTCCAGTCCAACTC	6338
Db	5941	GTTTGGGACTGGGTGTGCACGGTGTGAGTGACTTCAAGACTTGGCTTCCAGTCCAACTC	6000
Qy	6339	CTGCGCGGTTTACCGGAGTCCCTTTCCTGTATGCCAAACGCGGGTACAGGGAGTCTGG	6398
Db	6001	CTGCGCGGTTTACCGGGTGTCCCTTTCCTCTCATGCCAAAGTGGGTACAAGGGAGTCTGG	6060
Qy	6399	CGGGGGACGCATCATGCAAAACCACTGCCATGCGGAGCACAGATCGCGGACATGTC	6458
Db	6061	CGGGGGACGCATCATGCACCACTGCCATGTGGAGCACAGATCGCGGACATGTC	6120
Qy	6459	AAAAACGGTTCCATGAGGATCGTAGGGCTAGAACCTTCGACGCAACACGTGGCACGGAACG	6518
Db	6121	AAAAACGGTTCCATGAGGATCATCGGGCGGAAACCTTCGACGCAACACGTGGCATGGAACA	6180
Qy	6519	TTTCCCATCAACGCATACACAACCGGACCTTTGCAACACCTCCCGGCGCCCACTATTCC	6578
Db	6181	TTTCCCATCAACGGGTACACACCGGCGCCCTTCGACGCTTCCCGGCGCCCAACTATTCC	6240
Qy	6579	AGGGCGCTATGGCGGTGGCTGCTGAGGAGTACGTGGAGGTACGCTGTGGGGGATTTTC	6638
Db	6241	AAGGCGCTGTGGCGGTGGCTGCTGAGGAGTACGTGGAGGTTCACGCGGGTGGGGATTTTC	6300

QY 6639 CACTACGTGACGGGATGACCACTGACAAACGTAAGTGCCTATGCCAGTGTCCGGCCCC 6698
Db 6301 CACTACGTGACGGGATGACCACTGACAAACGTAAGTGCCTATGCCAGTGTCCGGCCCC 6360
QY 6699 GAATTCCTTACGGAGGTGATGAGTGGGTGTCACAGGTAGCTCCGGCGTGCAAACTT 6758
Db 6361 GAGTTTTCACGGAGGTGATGAGTGGGTGTCACAGGTAGCTCCGGCGTGCAAACTT 6420
QY 6759 CTTCACGGAGGACGTTCACGTTCCAGGTCGGGCTCAACCAATATCTTGGTCGGGTCCGAG 6818
Db 6421 CTCCTACGGAGTGAAGTTGATTTACAGTCCGGGCTCAATCAATACCTGTTGGGTACAG 6480
QY 6819 CTCCTACGGAGTGAAGTTGATTTACAGTCCGGGCTCAATCAATACCTGTTGGGTACAG 6878
Db 6481 CTCCTACGGAGTGAAGTTGATTTACAGTCCGGGCTCAATCAATACCTGTTGGGTACAG 6540
QY 6879 CACTACGTGACGGGATGACCACTGACAAACGTAAGTGCCTATGCCAGTGTCCGGCCCC 6938
Db 6541 CACTATACAGCAGAGGGGCTAAGCGTGAAGTGGGTCAGGGGCTCTCCGCCCTCTTGGCC 6600
QY 6939 AGCTCATCAGCTAGCCAGTGTCTGGCCCTCTTGAAGGCGACATGCACTACCC- --- 6993
Db 6601 AGCTCTTACAGCTAGCCAGTGTCTGGCCCTCTTGAAGGCGACATGCACTACCCATCT 6660
QY 6994 -----ACCATGACTCCCGGACGCTGACCTCATCGAGGCCAACTCTTGTGGCGGCGAG 7046
Db 6661 TCCATATCTTGACTCTCCGGAGCTCGACCTCATTTGGGCCAACTCTCTGTGGCGGCGAG 6720
QY 7047 GAGATGGCGGAAAATCATCTCGGTGGAGTCAGAGAATAAGGTAGTAATTTCTGACTCT 7106
Db 6721 GAGATGGCGGAAAATCATCACCGCGTGGAGTCGGAGAACAGGTGGTAGTCTCTAGACTCT 6780
QY 7107 TTTCGAACCGCTTCACGCGAGGGGATGAGAGGAGATATCGTCGGCGGAGATCCCTG 7166
Db 6781 TTTCGAGCGCTTCGAGCGAGGGGATGAGATGAATAATATTCATTGCGCGGAGATCCCTG 6840
QY 7167 CGAAATCCAGGAAGTTCCCTCAGCGTTGCCATATGGGCACGCGGACTACAACTCT 7226
Db 6841 CGGAAGTCCAAGAGTTCCCGGGCGATACCCATATGGGCGCGGATTTACAACTCT 6900
QY 7227 CCACTGCTAGATCTCGGAAGACCGGACTACGTCCTCGGTGGTACACGGATGCCCA 7286
Db 6901 CCAATTGTTAGAGTCTTGGAAGAACCGGACTACGTCCTCGGTGGTACACGGGTGCCCA 6960
QY 7287 TTGCGACCTTACCAAGCTCTCCAAATACCACTCCACGAGAAAGAGGAGGTTCTCTG 7346
Db 6961 TTGCGACCTGTCAAGGCGCTTCAATACCACTCCACGAGAAAGAGGAGGTTCTCTG 7020
QY 7347 ACAGAAATCCAAATGTGTCTCTGCTTTGGCGGAGCTCGCCACTAAGACCTTCGCTAGCTCC 7406
Db 7021 ACGGACTCCACCGTGTCTCTGTTTGGCGGAGTCTGCTACCAAAACCTTCGGCGAGCTCC 7080
QY 7407 GGATCGTCGGCGGTGATAGCGGACCGGACCGGCTTCTGACTGCGCTCCGACGAC 7466
Db 7081 GAATGTGCGCGCGCGACACGCGGACCGGCGCGGCTCTGACACAGACCTCCGACAAAC 7140
QY 7467 GGTGACAAAGGATCCGACGTTGAGTCGTACTCTCTCATGCCCCCTTCAAGGGGAGCGG 7526
Db 7141 GCGGCAAAAGATCTCGACGCTGAGTCATGCTCTCTATCCCCCTTGGAGGGGAGCGG 7200
QY 7527 GGGGACCCCGATCTCAGCGACGGGCTTGTGTCTACCGTGAGTGAGGAGGCTAGTAGGAT 7586
Db 7201 GGGGACCCCGATCTCAGCGACGGGCTTGTGTCTACCGTGAGCGAGAGGCTGTGTGAGC 7260
QY 7587 GTCGTGTGCTCAATGTCTATACGTGGACAGGCGGCTGATCAGCCATGCGCTGCG 7646
Db 7261 GTCGTGTGCTCAATGTCTATACATGACAGGTGCGCTGATCAGCCATGCGCGCGG 7320
QY 7647 GAGGAAGTAAGTGCCTCAACCGTTGAGCACTCTTGTGCTGCTCACCACACATG 7706
Db 7321 GAAGAAAGCAAGTGCCTCAACCGTTGAGCACTCTTGTGCTGCGCCATCACAAACATG 7380
QY 7707 GTCTACGCCACAACATCCCGCGCGCAAGCTCCCGGAGAAAGGTCACCTTTTGACAGA 7766

Db 7381 GTCTACGCCACGACATCCCGCAGCGCGCTCGCGCAGAAAGGTCACCTTTTGACAGA 7440
QY 7767 TTGCAAGTCTCGGATGATCATTACCGGACCTACTCAAGGAGATGAAGCGGAGCGGTCC 7826
Db 7441 CTGCAAGTCTCGGATGACCAATTACCGGACGTGCTTAAGGAGATGAAGCAAGGCGTCC 7500
QY 7827 ACAGTAAAGGCTAAGCTTCTATCTATAGAGGAGGCTGCAAGCTGACGCCCCCACAATTCCG 7886
Db 7501 ACAGTCAAGGCTAACTTCTATCCATAGAGAGCTGCGGCTGACGCCCCCACAATTCCG 7560
QY 7887 GCCAAATCCAAATTTGGCTATGGGCAAGGACGTCGGAACCTATCCAGAGGCGCGTT 7946
Db 7561 GCCAAATCCAAATTTGGCTATGGGCAAGGACGTCGGAACCTATCCAGCAGGCGCATC 7620
QY 7947 AACCATATCCGCTCGGTGGGAGACTTGTGGAGAGACTGTGACCAATTTGACACC 8006
Db 7621 AACCATATCCGCTCGGTGGGAGACTTGTGGAGAGACTGTGACCAATTTGACACC 7680
QY 8007 ACCATATGGCAAAAGTGAGGTTTTCTGCGTCCAAACAGAGAGGAGCGCCGCAAGCCA 8066
Db 7681 ACCGTATGGCAAAAGTGAGGTTTTCTGCGTCCAAACAGAGAGGAGCGCCGCAAGCCA 7740
QY 8067 GCTCGCTTATCGTATTTCCAGACCTGGAGTTCGTATGCGAGAGATGCCCTTTAC 8126
Db 7741 GCGCGCTTATCGTATTTCCAGATTTGGGAGTTCTGTATGCGAGAGATGCTCTCTAC 7800
QY 8127 GAGTGGTCTCCACCTTCTCGGCGCTGATGGCTCTCATAGGATTTCAATCTTC 8186
Db 7801 GATGTTGTTCTCACCTTCTCAAGCGGTGATGGCTCTCTCATCGGATTCAGTACTCT 7860
QY 8187 CCCAAGCAGCGGTGAGTTCCTGCTGAATACCTCGAAATCAAAGAAATGCCCTATGGC 8246
Db 7861 CCCGCGAGCGGTGAGTTCCTGTTAAAGCTGGAATCAAAGAAATCCCTATGGC 7920
QY 8247 TTCTCATATGACACCGCTGTTTGAATCAACGGTCACTGAGAGTGACATTCGTGTTAG 8306
Db 7921 TTCTCATATGACACCGCTGTTTGAATCAACGGTCACTGAGAAATGACATCGTGTGAG 7980
QY 8307 GAGTCAATTTACCAATTTGATGACTTTGGCCCCGAGGCGCAGACAGGCCATAGGTGCTC 8366
Db 7981 GAGTCAATTTACCAATTTGATGACTTTGGCCCCGAGGCGCAGACAGGCTATATAATCGCTC 8040
QY 8367 ACAGAGCGGCTTATCATCGGGGTCCTGACTAATCAAAAGGCGAGAACTGCGGTAT 8426
Db 8041 ACAGAGCGGCTTATATCGGGGTCCTGACTAATCAAAGGCGAGAGCTGTGGTTAT 8100
QY 8427 CGCGGTCGCGCGCAAGTGGCGTGTGACGACTAGCTGCGGTAATACCTCAATGTTAC 8486
Db 8101 CGCGGTCGCGCGAGCGGCTGTGACGACTAGCTGCGGTAATACCTCAATGTTAC 8160
QY 8487 TTGAAGGCCATGTCAGACCTGTGAGCTGCAAGCTCCAGGACTGACAGTGTCTGTAAC 8546
Db 8161 TTGAAGGCCCTGTCGCGCTGTGAGCTGCAAGCTCCAGGACTGCAAGTGTCTGTAAC 8220
QY 8547 GGAGACGACTTGTGTTATCTGTAAGCGGGGAAACCCAGAGGATGCGGCGGCTTA 8606
Db 8221 GGGGACGACTTGTGTTATCTGTAAGCGGGGAAACCCAGAGGATGCGGCGGCTTA 8280
QY 8607 CGAGCTTTCAGGAGGCTATGACTAGGTTATCCGCCCCCCCCGGGATCCGCCCCCAACA 8666
Db 8281 CGAGCTTTCAGGAGGCTATGACTAGGTTATCCGCCCCCCCCGGGACTTGGCCCCAACA 8340
QY 8667 GAATACGACTGGAGCTGATTAACATGTTCTTCAATGTGTGAGTGGGCAAGTGA 8726
Db 8341 GAATACGACTGGAGTGTGATTAACATCATGTTCTTCAATGTGTGAGTGGGCAAGTGA 8400
QY 8727 TCTGCAAAAGGCTATGACTTACCTCACCGGTACCCACACCCCTTTCAGGGGTGCG 8786
Db 8401 TCTGCAAAAGGCTGTACTTACTCTGCGATCCACACCCCTTTCAGGGGTGCG 8460
QY 8787 TGGGAGACGCTTAGACACATCTCAATCAACTCTTGGCTAGGCAATATCATGTATGCG 8846

Db 8461 TGGGAACAGCTAGACACACCTCCAGTTAACTCTCTGGCTAGGCAACATATATCATGTATGG 8520
Qy 8847 CCACACCTATGGCAAGGATGATTTCTGATGACTCACTTTTCTCCATCTCTTAGCTCAA 8906
Db 8521 CCACACCTATGGCAAGGATGATTTCTGATGACCACTTTCTTCCATCTCTTAGCTCAG 8580
Qy 8907 GAGCAACTTGAAGAAAGCCCTGGATTTGTAGATCTACGGGCTTTGCTACTCCATTTAGGCA 8966
Db 8581 GAGCAACTTGAAGAAAGCCCTGGATTTGCCAAATCTACGGGCTTTGCTACTCCATTTAGGCA 8640
Qy 8967 CTTGACCTACTCAGATCATATGAACAGCTCCATGCTTTAGCGCATTTACACTCCACGT 9026
Db 8641 CTTGACCTACTCAGATCATATGAAGAGCTCCATGGTCTTAGCGCATTTTCACTCCATAGT 8700
Qy 9027 TACTCTCCAGGTGAGATCAATAGGTTGGCTTCTATGCTCAGGAACCTTTGGGGTACCACCC 9086
Db 8701 TACTCTCCAGGTGAGATCAATAGGTTGGCTTCTATGCTCAGGAACCTTTGGGGTACCACCC 8760
Qy 9087 TTGCGAACCTTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACTGTCCCGAGGGGGG 9146
Db 8761 TTGCGAGTCTGGAGACATCGGGCCAGGAGCTCCGCGCTAACTACTGTCCCGAGGGGGG 8820
Qy 9147 AGGCGCCCACTTTGTGGCAGATACCTCTTTAACTGGGCACTAAGGACCAAGCTTAAATC 9206
Db 8821 AGGCGCCCACTTTGGCGCAATACCTCTTCACTGGGCACTAAGGACCAAGCTTAAATC 8880
Qy 9207 ACTCCAAATCCCGCGCTCCAGCTGGACTTTGTCTGGCTGGTTGCTGCTGTTACAGC 9266
Db 8881 ACTCCAAATCCCGCGCTCCAGCTGGACTTTATCCGCTGTTGTTGCTGCTACAGC 8940
Qy 9267 GGGGAGACATATATACAGCTGTTCTGTCGCCGACCCCGCTGTTTCCGTTGTCCTA 9326
Db 8941 GGGGAGACATATATACAGCTGTTCTGTCGCCGACCCCGCTGTTTCCGTTGTCCTA 9000
Qy 9327 CTCCTACTTTCTGAGGGTAGGCAATTAACCTGCTCCCAACCGATGAACGGGAGCTAA 9386
Db 9001 CTCCTACTTTCTGAGGGTAGGCAATTAACCTGCTCCCAACCGATGAACGGGAGCTAA 9060
Qy 9387 CCACCTCAGGCC-TTAAGCCATTTCCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9445
Db 9061 ACACCTCAGGCCAATAGGCCAATTTCCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9120
Qy 9446 TTTTTTTTCTTTCT 9500
Db 9121 TT 9180
Qy 9501 GGCTCCATCTTAGCCCTAGTCAAGGCTAGCTGTGAAGGTCGGTAGCGCATGACTGCA 9560
Db 9181 GGCTCCGCTTTAGCCCTAGTCAAGGCTAGCTGTGAAGGTCGGTAGCGCATGACTGCA 9240
Qy 9561 GAGAGTGTGATAGTGGCTCTCTGACATCATGT 9595
Db 9241 GAGAGTGTGATAGTGGCTCTCTGACATCATGT 9275

RESULT 8
US-10-189-359-13
; Sequence 13, Application US/10189359
; Publication No. US20040039187A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, Annette
; APPLICANT: SANGAR, DAVID V.
; APPLICANT: LEMON, STANLEY M.
; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 9599

; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(9377)
US-10-189-359-13

Query Match 66.0%; Score 6335.8; DB 7; Length 9599;
Best Local Similarity 78.9%; Pred. No. 0;
Matches 7573; Conservative 0; Mismatches 2022; Indels 4; Gaps 2;

Qy 1 GCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCTCTGTGAGGAATCTACTG 60
Db 1 GCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCTCTGTGAGGAATCTACTG 60
Qy 61 TCTTACGACAGAAAGCGTCTAGCCATGGGGTTAGTATGATGTGCTGCGAGCTCCAGGAC 120
Db 61 TCTTACGACAGAAAGCGTCTAGCCATGGGGTTAGTATGATGTGCTGCGAGCTCCAGGAC 120
Qy 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTTGCCAG 180
Db 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTTGCCAG 180
Qy 181 GACGACGGGTCTCTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTGCCTCC 240
Db 181 GACGACGGGTCTCTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTGCCTCC 240
Qy 241 GCAGACTGTACCGAGTAGTCTTGGTCCGAAAGGCTTGTGTACTGCTCTGATAGG 300
Db 241 GCAAGACTGTACCGAGTAGTCTTGGTCCGAAAGGCTTGTGTACTGCTCTGATAGG 300
Qy 301 GTCTTCCGAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTTAAAC 360
Db 301 GTCTTCCGAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTTAAAC 360
Qy 361 CTCAAGAAAAAACAACGCTTAACCAACCCGCCCAACAGGACGTCAAGTTCCCGGGCG 420
Db 361 CTCAAGAAAAAACAACGCTTAACCAACCCGCCCAACAGGACGTCAAGTTCCCGGGCG 420
Qy 421 GTGTTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCGAGGGGCCCGAGTTGGGTGTC 480
Db 421 GCGGTTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCGAGGGGCCCGTAGATTGGGTGTC 480
Qy 481 GCGGCTAGAGAGGCTTCGAGCGGTTCGCAACCTCTGGAAGGCGACACCTATCCAA 540
Db 481 GCGGCTAGAGAGGCTTCGAGCGGTTCGCAACCTCTGGAAGGCGACACCTATCCCA 540
Qy 541 AGGCTCCCGACCCGAGGCGAGGCGCTGAGGCTCAGCCGCTACCTTTGGCCCTCTATG 600
Db 541 AGGCACTCCGCGCGAGGCGAGGCGCTGAGGCTCAGCCGCTACCTTTGGCCCTCTATG 600
Qy 601 GCAATGAGGGCTTGGGCTGGGAGGAGTGGCTCTCTGTCAACCCCGGGCTCCCGGCTAGTT 660
Db 601 GCAATGAGGGTTGGGCTGGGAGGAGTGGCTCTCTGTCTCCCCGTTGGCTCTCGGCTAGCT 660
Qy 661 GGGGCCCCACGACCCCGGGCGTAGGTGCGCAATTTGGGTAAAGTATCATGATACCTTTA 720
Db 661 GGGGCCCCACAGACCCCGGGCGTAGGTGCGCAATTTGGGTAAAGTATCATGATACCTTTA 720
Qy 721 CATCGGCTTCGCGATCTCATGGGTACATTCGCTCGTGGCGGCCCTCTAGGGGCG 780
Db 721 CGTGGGCTTCGCGACCTCATGGGTACATTCGCTCGTGGCGGCCCTCTTAGAGGCG 780
Qy 781 CTGCCAGGCTTGGCAGCGGTTCGGGTTCGGAGGAGCGGCTGAACATATGCAACAG 840
Db 781 CTGCCAGGCTTGGCAGCGGTTCGGGTTCGGAGGAGCGGCTGAACATATGCAACAG 840
Qy 841 GGAATTCGCGGTTCCTTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 841 GGAATTCCTGGTTCCTTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Qy 901 TCCAGCTTCCGCTTATGAAGTGCACACGCTGTCGGGATATACCATGTACGAAACGACT 960
Db 901 TCCAGCTTCCGCTTATGAAGTGCACACGCTGTCGGGATATACCATGTACGAAACGACT 960

Db 901 TGGCCGCTTACGCTACCAAGTCGCAATTCTCTCGGGCTTTTACCATGTCACCAATGATT 960
Qy 961 GCTCCAACTCAAGCAATGTGTATGAGGACAGCGGACGTGATCATGATCACTCCCGGGTGG 1020
Db 961 GGCCTAACTCGAGTATTTGTATACGAGCGCGCGATGCCATCTCTGCACACTCCCGGGTGTG 1020
Qy 1021 TGCCTGTGTTTACGAGGGTAAACAGCTCCGTTGCTGGGTAGGCTCACTCCACGCTCG 1080
Db 1021 TCCCTTGGCGTTTCGAGAGGTAAACGCTCGAGGTGTGGGTGGCGGTGACCCCAACGGTGG 1080
Qy 1081 CGGCCAGGAATGCCAGCGTCCCACTTACGACAATACGACGCCACGTCGACTTGTCTGTTG 1140
Db 1081 CCACACAGGACGGCAAACTCCCAACACGACGCTTCGACGTATATCGATCTGCTTGTGCG 1140
Qy 1141 GGAAGGCTGCTTCTGCTCCGCTATGTAGCTGGGGATCTCTGCGGATCTATTTTCCCTCG 1200
Db 1141 GGAGCGCAACCTCTGCTCGGCCCTCTACGTGGGGACCTGTGCGGGTCTGCTTCTTCTTG 1200
Qy 1201 TCTCCAGCTGTTCACCTTCTCGCTCGCGCGCATGAGACAGTGCAGGACTGCAACTGCT 1260
Db 1201 TTGGTCAACTGTTTACCTTCTCTCCAGCGCCACTGGACGACGCAAGACTGCAATTTGTT 1260
Qy 1261 CAATCTATCCCGGCATGATACAGTTCACGTCACCGATGCTTGGGATATGATGAACTGCT 1320
Db 1261 CTATCTATCCCGGCATATACCGGTATCGCATGSCATGGGATATGATGAACTGCT 1320
Qy 1321 CACTTACAAACAGCCCTAGTGTGTGCGAGTTGCTCGGATCCCAAGCTGTCTGTGGACA 1380
Db 1321 CCCCTACGCGAGCTTGGTGGTAGTCTCAGCTGCTCGGATCCCAAGCCATCATGGACA 1380
Qy 1381 TGGTGGCGGGGCCACCTGGGAGTCTCGGGGCGCTTGCCTACTATTCCATGGTAGGGA 1440
Db 1381 TGATCGCTGTGCTCACTGGGGAGTCTCGCGGGCATAGCTGATTTCTCCATGGTGGGA 1440
Qy 1441 ACTGGCTAAGTCTGATTTGGCGCTACTCTTTGCGCGCTTGCAGGGGAGACCCACA 1500
Db 1441 ACTGGCGAAGTCTGTTGGTAGTGTGCTGCTATTTGCGGGCTGCAAGCGGAAACCCACG 1500
Qy 1501 CGACGGGAGGGTGGCGGCGCACACACCTCGGGGTTTACGTCCTTTTCTCATCTGGGG 1560
Db 1501 TCACCGGGGAAATGCCGCGCGCACACACCGCTGGGCTTGTGGTCTCTTACACAGGG 1560
Qy 1561 CGTCTCAGAAATTCAGCTTTGTAATACAAACGAGCTGGACATCAACAGGACTGCC 1620
Db 1561 CCAAGCAGAACTCAAACTGATCAACACCAACGCGAGTTGGACATCAATACACAGCGCT 1620
Qy 1621 TAAATGCAATGACTCCCTCCAACTGGGTTCTTTGCGCGCTGTTTACGACACAAAT 1680
Db 1621 TGAATGCAATGAAAGCTTAAACCGGCTGGTTAGCAGGGCTCTTCTATCAACACAAAT 1680
Qy 1681 TCAACTCGTCCGGGTGCCCGGAGCGATGGCGAGCTGCCGCCCAATTGACTGGTTCCGCC 1740
Db 1681 TCAACTCTTCAGGCTGCTTGAGAGGTTGGCCAGCTGCCGACGCTTACCGATTGTC 1740
Qy 1741 AGGGTGGGGCCCCATCACTATATCAAGCTTAACAGCTCGGATCAGAGCCCTTATGCT 1800
Db 1741 AGGGCTGGGGTCTATCAGTATGCAACGGAAGCGGCTCGACGAAACCCCTTACTGCT 1800
Qy 1801 GGCATTAACGGCTCGACCGTGTGTGCTGTACCGCGCTGCGAGGTGTGTGTGTCAGTGT 1860
Db 1801 GGCATTAACCGCTTCCAGACCTTGTGGCAATTTGTCCCGCAAGAGCGGTGTGTGCGCGGTAT 1860
Qy 1861 ATTGTTTACCCCAAGCCCTGTTGTGGTGGGACACACCATGCTTTCGGGTGCTTACGT 1920
Db 1861 ATTGTTTCACTCCAGCCCGTGGTGGTGGGAACCGACAGGTTCGGGCGGCTTACTCT 1920
Qy 1921 ATAGCTGGGGGAGATGAGACAGCGTATGCTCTCTCAACAAACAGCGCTCCGCCACAAG 1980
Db 1921 ACAGCTGGGGTGCAAAATGATACGGATGCTCTCGTCTTAAACAAACACGAGCCACCGCTGG 1980
Qy 1981 GCAATGGTTCGGCTGTACATGGATGAATAGTACTGGGTTCACTAAGAGCTGCGGAGTTC 2040
Db 1981 GCAATTGGTTCGGTGTGATCCTGGATGAATCAACTGGATTACCAAAAGTGTGCGGAGGCG 2040

Qy 2041 CCCCGTGTAAATCATCGGGGGGTCCGTAAACCGACCTTGATCTGCCCCACGGAAGTCTGCTTC 2100
Db 2041 CCCCTTGTGTATCATCGGAGGGTGGGCAACAAACACCTTGTCTGTGCCCCACTGATTGCTTC 2100
Qy 2101 GGAAGCACCCGAGGCTACTTACACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT 2160
Db 2101 GAAAACATCCGGAAGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATTACACCAAGT 2160
Qy 2161 GCCTAGTAGATACCCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTTCATCT 2220
Db 2161 GCATGTCGATCTACCCGTATAGGCTTTGGCACTATCTTGTACCATCAATTACACCATAT 2220
Qy 2221 TTAAGTTAGATGATGTGGGGGCGTGGAGACAGAGCTCAATCCGCGATCAATTTGA 2280
Db 2221 TCMAAGTCAGGATGTACGTGGAGGGGTCCGAGCAGAGCTGGAGCGGCTGCAACTGGA 2280
Qy 2281 CTCGAGGAGAGCGCTGTAACTTTGGAGGACAGGATAGTTCAGAACTCAGCCCGCTGCTG 2340
Db 2281 CGCGGGCGAAACGCTGTGATCTGGAAGACAGGACAGGTCCGAGCTCAGCCGTTGCTGCG 2340
Qy 2341 TGTCTACAAACAGAGTGGCAGATACCTGCCCTGTGCTTTTACACACCTTACCGGCTTATCCA 2400
Db 2341 TGTCCACACACAGTGGGAGGTCCTTCCGTGTTCTTTACGACCTTGCAGGCTTGTCCA 2400
Qy 2401 CTGGTTTGCATCCATCTCOATCAGAACTCGTGGACGTGCAATACCTGTACGGTGTAGGT 2460
Db 2401 CGGGCTCATCCACCTCCACAGAACTTGTGGACGTGCAGTACTGTACGGGTAGGCT 2460
Qy 2461 CAGCGTTTGTCTCTTTTGCATCAATGGGAGTACATCTCTGTGTGCTTTTCTTCTCTCTG 2520
Db 2461 CAAGCATCCGCTCTCGGGCCATTAAAGTGGGAGTACGTCGTTCTCTGTCTTCTCTCTG 2520
Qy 2521 CAGACCGCGGTGTGCTGCTTGTGGATGATCTGCTGATAGCCAGGCTGAGGCG 2580
Db 2521 CAGACCGCGGTGTGCTGCTTGTGGATGATGTTACTCATATCCCAAGCGGAGGCG 2580
Qy 2581 CTTAGAGAACTTGGTGTCTCTCAATGGCGGCTCGTGGCCGAGCGCATGTTATTTCT 2640
Db 2581 CTTTGGAGAACTCTGTAATACTCAATGACGATCCTCGCGGGACGACGCTTGTGT 2640
Qy 2641 CTTTCTTGTGTTCTTCTGCGCGCTGTATCAATTAAGGCGAGGCTGCTCTGCGGCG 2700
Db 2641 CTTCTCTGTGTTCTTCTGCTTTGCGTGTATCTGAAGGGTAGGTGGGTGCGCGGAGCG 2700
Qy 2701 CGTATGCTTTTATGGCGTATGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Db 2701 TCTACGCCCTCTACGGGATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Qy 2761 CTTAGCGCTTGGACCGGGAGATGGCTGCATCGTGGGGGTGCGGTTCTTGTAGGTCTCG 2820
Db 2761 CATACGCACTGGACACGAGGTTGGCGCGCTCGTGTGGCGGCTTGTCTTGTTCGGGTTAA 2820
Qy 2821 TATTCCTGACCTTGTGCACATCAATAAAGTGTCTCTCATAGGCTCATATGGTGGTTAC 2880
Db 2821 TGGCGCTGACTCTGTGCCATATTAACAAGCGCTATATCAGCTGGTGCATGTGTGGCTTC 2880
Qy 2881 AATACTTTTATCACCAGACCGAGCGCATGCAAGTGGGTGCGGCTCCCGCTCAAGGTTTC 2940
Db 2881 AGTATTTTCTACCCAGATAGAACGCACTGCAGTGTGGGTTCCCGCTTCAACGCTCC 2940
Qy 2941 GGGGAGCGCGGATGCCATCATCTCTCTCACTGTGCGGTTTCATCCAGAGTTTAAATTTTG 3000
Db 2941 GGGGGGGCGGATGCCGCTCATCTTACTCATGTGTGTAGTACACCGACCTTGGTATTG 3000
Qy 3001 ACATCAACAACTCTGTCTGCGCATATCTGGCGCGCTCATGTGTCTCAGGCTGGCATAA 3060
Db 3001 ACATCAACAACTACTCTGCGCATCTTCGGACCTTTCGATTTCTTCAAGCCAGTTTCG 3060
Qy 3061 CGAGAGTCCGCTACTTCTGTCGCGCTCAAGGCTCATTCGTGCATGCTGTTAGTTCGAA 3120
Db 3061 TTAAGTCCCTTACTTCTGTGCGGCTTCAAGGCTTCTCCGATCTGCGGCTAGCGCGA 3120

QY	3121	AAAGTCGCGGGGTCAATATATGTCCAAATGGTCTTCATGAAGCTGGCGCGCTGACAGGTA	3180
Db	3121	AGATAGCCGGAGTCAATACGTGCAAAATGSCCATCATCAAGTTAGGGCGCTTACTGGCA	3180
QY	3181	CGTAGCTTTATACCAATCTTACCCCACTGGGGACTGGGCCCAACGGGGCCTTACAGACC	3240
Db	3181	CCTATGTGTATACCAATCTCACCCCTTTCGAGACTGGGCGCACAAACGGCCTGCCAGATC	3240
QY	3241	TTCCGGTGGCGGTAGAGCCGCTCGTCTCTCCGCCATGGAGACCAAGGTCATCACTCGG	3300
Db	3241	TGGCCGTGGCTGTGGAAACAGTCTCTTCTCCGAATGGAGACCAAGCTCATCACTGGG	3300
QY	3301	GAGCAGACCCGCTGCGGTGTGGGACATCATCTTGGGTCTACCCGTCTCCGCCGAAGGG	3360
Db	3301	GGGCAGATACCGCGCGTGGCGTGACATCATCAACGGCTTGCCCGTCTCTGCGCGTAGGG	3360
QY	3361	GGAAAGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTGGCGACTCCTTG	3420
Db	3361	GCCAGGAGATCTGCTTGGGCCAGCCGACCGGAATGGTCTCCAAGGGGTGGAGGTTGCTGG	3420
QY	3421	CGCCCATCAGGCGCTACTCCCAACAAACGGGGCGTACTTGGTTGCATCATCACTAGCC	3480
Db	3421	CGCCCATCAGGCGGTAGCCCAACAGAGAGGCCCTCTAGGGTGTATATCAACAGCC	3480
QY	3481	TCACAGCCCGGACAAAGAACAGGTCCAAAGTGGGAGGTTCAAGTGGTTTCTACCGCAACAC	3540
Db	3481	TGACTGGCCGGACAAAACCAAGTGGAGGTCAGATCGTGTCAACTGTCTACCC	3540
QY	3541	AATCTTCTCGGGAACCTTGATCAACGGCGTGTCTGGACTGTCTACCATGGCGCTGGCT	3600
Db	3541	AAACCTTCTCGGCAACGTGATCAATGGGGTATGCTTGGACTGTCTACCAACGGGGCCGAA	3600
QY	3601	CGAAGACCTTAGCGGTGCAAAAGTCCAAATACCCAAATGTACACAAATGTAGACCTCG	3660
Db	3601	CGAGGACCATCGATCAACCCAGGGTCTCTGTATCCAGATGTATATCAATGTGACCAAG	3660
QY	3661	ACCTCGTGGCTGGCAGGCGCCCCCGGGCGGCTCCATGACACCATGAGCTGTGGCA	3720
Db	3661	ACCTTGTGGCTGGCGCGCTCCTCAAGGTTCCCGCTCAITTGACACCGCTGTACTCGGCT	3720
QY	3721	GCTCGGACTTTACTTTGGTCAAGAGACATGCTGATGTCTCATTCGGTGGCGCGGAGGCG	3780
Db	3721	CCTCGSACCTTTACTTGGTCAAGAGGCAACCCGATGTCAITTCGGTGGCGCGGAGGCTG	3780
QY	3781	ACAGCGGGAAGTCTACTCTCCCGCAGGCGCGCTCTCTACTCTGAAGAGGCTCTCGGGTG	3840
Db	3781	ATAGCGGGTAGCTGCTGTTTGGCCCGGCGCCATTTCTTACTTGTGAAGGCTCTCTCGGGG	3840
QY	3841	GTCCATTTGCTTTGCCCTTCGGGGACGTCGTGGGCGTCTTCGGGGTGTGTGTGCAACC	3900
Db	3841	GTCCGCTGTTGTGCCCCCGGGACACGCCGTGGGCTTATTCAGGCGCGGTTGTGCAACC	3900
QY	3901	GGGGGTCGGAAGCGGTGGACTTCAATACCGGTTGAGTCTATGGAAACTACCATGGGT	3960
Db	3901	GTGGAGTGGCTTAAACGCGTGGACTTTATCCCTGTGGAGAACCTTAGGACAACCATGAGAT	3960
QY	3961	CTCCGCTCTTCACAGACAACCTCAACCCCCCGGCTGTACCGCAGACAATCCAAAGTGGCAC	4020
Db	3961	CCCCGGTGTTCACGGACAATCTCTCTCAACAGCAGTGCCTCCAGAGCTTCCAGGTGGCCC	4020
QY	4021	ATCTGCAGCTCTACTTGGCAGCGGCAAGAGCACAAAGTGCCTGGCTGTGATGCAAGCCC	4080
Db	4021	ACCTGCATGCTCCCAACCGCAGCGGTAAAGACACCAAGTCCCGGCTGCGTACGACGCC	4080
QY	4081	AAGGTTACAAGGTGCTGTCGTGAACCGGTCCGTTGGCGCACTTATAGGGTTTGGGGCGT	4140
Db	4081	AGGGCTACAAGGTGTGGTGTCTAACCCCTCTGTGTGTCGAACGCTGGGCTTTGGTGCTT	4140
QY	4141	ATATGTCCAAGGCAACGGTATCGACCTCAATCAATCAGAACTGGGGTAAGGACCATTACCA	4200
Db	4141	ACATGTCCAAGGCCCATGGGGTGTATCTCTATATACGACCGGGGTGAGAACAAATTACCA	4200
QY	4201	CGGGCGGCTCCATTTACGTACTCCACCTATGCGAAGTTCCTTGGCCGACGGTGGCTGTCTG	4260

Db	4201	CTGGCAGCCCACTCAGTACTCACCTACGGCAAGTTCCTTGGCCGAGCGGGTGTCTCAG	4260
Qy	4261	GGGGCGCTATGACATCATAAATATGTGATGAGTGCACATCAACTGACTCGACTACCATCT	4320
Db	4261	GAGTGTCTATGACATAATATATTTGTGACGAGTGCACATCCACGGATGCCACATCCATCT	4320
Qy	4321	TGGGCATCGGCACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGCGCTCGTCTGTCTCG	4380
Db	4321	TGGGCATCGGCACGTCTCTTGTGACCAAGCAGAGACTCGGGGGCGGAGACTGGTGTGTCTCG	4380
Qy	4381	CCACCGCTACACCTCCGGGATCGGTTTACGTCGCCACACCCCAATATTCGAGGAAATAGGCC	4440
Db	4381	CCACTGTCTACCCCTCCGGGCTCCGTCACTGTGTCCATCTCTAATCTGAGGAGTTGCTC	4440
Qy	4441	TGTTCCAAATGAGAGATCCCCCTTCTATGGCAAGCCATCCCCATTTGAGGCCCATCAAGG	4500
Db	4441	TGTTCCACCACCGAGAGATCCCCCTTTACGGCAAGGCTATCCCCCTCGAGTGATCAAGG	4500
Qy	4501	GGGGAGGCATCTCATTTTCTGTCATTTCCAAAGAAATGTGACGAGCTCGCCGCAAGC	4560
Db	4501	GGGGAAGACATCTCATCTTCTGCGCATCTCAAGAGAGAGTGCAGCAGCTCGCCGCGAAGC	4560
Qy	4561	TGACAGCCTCGACTGAAAGCTGTAGCATATTTACCGGGCCCTTGATGTCTCGTCTATAC	4620
Db	4561	TGTCGCAATTTGGGCATCAATGCCGTGGCCCTACTACCGCGGCTCTTGACGCTGTCTATCC	4620
Qy	4621	CGCTATTCGGAGACGTCTGTGTCTGGTGCACACAGACGCTCTAATGACGGGTTTCAACGGCG	4680
Db	4621	CGACCGGCGGATGTGTCTGTCTGTCGACCGATGCTCTCATGACTGGCTTTACCGGGC	4680
Qy	4681	ATTTTGACTCAGTGATCGCATGTCAATATATGTGTACCCAGACAGTCTGACTTTCAGCTTGG	4740
Db	4681	ACTTCGACTCTGTGATAGACTGCAACACGCTGTGTCTACTCAGACAGTCTGATTTTCAGCCTTG	4740
Qy	4741	ATCCCACTTCAACATTGACAGCAGACCGTGCCTCCCAAGACGGGTCTCGCGCTCGCAAC	4800
Db	4741	ACCTCACTTTACATTTGACACACACGCTCCCCCAGGATGCTGTCTCTCAGGACTCAAC	4800
Qy	4801	GGCGAGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTTCCAGGAGAAC	4860
Db	4801	GCCGGGACGAGCTGGCAGGGGGAAGCCAGGCATCTATAGATTTGTGGCACCGGGGAGC	4860
Qy	4861	GGCCCTCGGGCATGTTTCGATCTTTCGGTCTGTGTGAGTGTATGACCGGGGCTGTGCTT	4920
Db	4861	GGCCCTCGGGCATGTTTCGACTCGTCTCGTCTCTGTGTGAGTGTATGACCGGGGCTGTGCTT	4920
Qy	4921	GGTATGAGCTACGCCCCCGAGACTTACAGTTAGGCTACGAGGCTATGNAACACCCGG	4980
Db	4921	GGTATGAGCTACGCCCCCGAGACTTACAGTTAGGCTACGAGGCTATGNAACACCCGG	4980
Qy	4981	GGTTGCCCGCTCTGCAGGACCATCTGGAGTTTCTGGGAGAGCGTCTTTCACAGGCTCAACC	5040
Db	4981	GGCTTCCCGTGTGCAGGACCATCTTGAAATTTTGGGAGGGCGTCTTTACGGGCTCACTC	5040
Qy	5041	ACATAGATGCCACTTCTCTGTCCAGACTAAACAGGCGAGAGACAACTTTCCTTACCTGG	5100
Db	5041	ATAAGATGCCACTTTTATCCCAACAAAGCAGAGTGGGGGAGAACTTCTCTTACCTGG	5100
Qy	5101	TGGCATATCAGCTTACAGTGTGGCCAGGGCTCAAGCTCCACCTCCATCTGTTGGGACCAAA	5160
Db	5101	TAGCGTACCAAGCACCGTGTGCGCTAGGGCTCAAGGCCCTTCCCCCATCTGTGGGACCAAG	5160
Qy	5161	TGTGGAGTGTCTCATACGGCTGAAACCTACACTGACGGGCCAACACCCCTGCTGTATA	5220
Db	5161	TGTGGAGTGTTTGATCCGCCCTTAAACCCACCTCCATGGGCCCAACACCCCTGCTATACA	5220
Qy	5221	GGCTAGGAGCGGTCCAAATAGAGTCTATCTTCAACACCCCATTAATTAATATCATCTGG	5280
Db	5221	GACTGGGCGCTGTTTCAAGTCAAGTCAACCTTGAGCACCCCAATCAACCAATATCATCATGA	5280
Qy	5281	CATGCATGTCCGCTGACCTGGAGGTCTGCTACCTAGCACCTGGTGTGTGTAGCGGAGTCC	5340

Db 5281 CATGATGTCGGCCGACCTGGAGTGCCTCAGAGACACCTGGGTGCTCGTTGGCGCGTC 5340
Qy 5341 TTGCAGCTTTGGCCGCATACCTGCCTGACGACAGGAGTGTGCTCACTTGTGGCGAGATCA 5400
Db 5341 TGGCTGCTCTGGCCGCGTATTGCTTGTCAACAGGCTGCGTGTCTATAGTGGCGAGATCG 5400
Qy 5401 TCTTGTCCGGGAAGCCAGCTGCTGTTCCCGACAGGGAAGTCTCTACAGGAGTTTCGATG 5460
Db 5401 TCTTGTCCGGGAAGCCGCAATTAACCTGACAGGAGGTTCTCTACAGGAGTTTCGATG 5460
Qy 5461 AGATGGAGAGTGTGCTCTCAACCTTCTTATCATCGAGGGAATGAGCTCGCCGAGC 5520
Db 5461 AGATGGAGAGTGTCTCTCAGCACTTACCGTATCATCGAGCAAGGATGATGCTCGCTGAGC 5520
Qy 5521 AATTCAAGCAAGAGCGCTCGGTTGTTGCAAAAGCCGACCAAGCAAGCGAGGCTGCTG 5580
Db 5521 AGTTCAAGCAGAGGCCCTCGGCCCTCTCGAGACCGCTCCGCCATCGAGAGGTTATCA 5580
Qy 5581 CTCCCGTGTGAGTCCAAAGTGGCGAGCCCTTGAGACCTTCTGGCGGAAGCACATGTGGA 5640
Db 5581 CCCCTGCTTCAGACCAACTGCGAGAACTCGAGGTCTTTTGGCGGAAGCACATGTGGA 5640
Qy 5641 ATTTCATCGCGGAATACAGTACTAGCAGGCTTATCCATCTGCTCTGGAACCCCGCGA 5700
Db 5641 ATTTCATCAGTGGGATACAAATCTTGGGGGCTGTCAACGCTGCTGTGTAACCCCGCA 5700
Qy 5701 TAGCATCATGATGCATTTACAGCTTCTATCAGTACCGCTACACCCCAACCAACACCC 5760
Db 5701 TTGCTTTCAATTGATGGCTTTTACAGTGGCGGTCAACAGCCCACTAACCACTGGCCAAACCC 5760
Qy 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGCTGCCCAACTCGCTCTCTCCAGCGCTGGT 5820
Db 5761 TCCTCTTCAACATATTGGGGGGTGGTGGCTGCCAGCTCGCCGCCCGCGCTG 5820
Qy 5821 CAGCTTTCGTGGCGCGGCATCGCCGAGCGGCTGTTGGCAGCATAGGCTTGGGAAG 5880
Db 5821 CTGCTTTGTGGGTGCTGGCTTAGTGGCGCGCCATCGGACAGCTGGACTGGGGAAG 5880
Qy 5881 TGCTGTGGACATCTGGCGGCTATGGGCAAGGGTAGCCGCGCATCTGTTGGCTTTA 5940
Db 5881 TCCTGTGGACATCTTGCAGGATATGGCGGGGTGGCGGAGCTCTTGTAGCATTA 5940
Qy 5941 AGGTATGAGCGCGAGTGCCTCCACCGAGACTGTGTCAACTTACTTCCCTGCGCATCC 6000
Db 5941 AGATCATGACCGTGANGTCCCTCCACGAGGACTGTGTCACTGTCTGCTGCCGCTCC 6000
Qy 6001 TCTCTCTGCTGCTGCTGCTGCTGGGTGCTGCGCAGCAATATCTGCTGCGCACGTGG 6060
Db 6001 TCTGCTGGAGCCCTTGTAGTCTGGTGTGGTCTGGCGAGCAATACTTGGCGCGCACGTTG 6060
Qy 6061 GCCCGGAGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCTTCCGTTCCGCGGTTA 6120
Db 6061 GCCCGGCGAGGGGGGAGTGGCAATGGAAGAACCGGCTAATAGCTTCCGCTCCCGGGGA 6120
Qy 6121 ACCAGCTCTCCCTACGCACTATGTGCTGTGAGCGGACGCTCAGCAGCTGTCACTCAGA 6180
Db 6121 ACCATGTTTCCCGCAGCACTAGTGTGCGGAGGAGTGCAGCGCGCTCACTGCCA 6180
Qy 6181 TCTCTCTAGCTTACCATCATCACTGCTGAAGCGGCTCACACGAGGATTAATAGG 6240
Db 6181 TACTCAGCAGCTCACTGTAAACAGCTCTGAGCGACTGCATCAGTGGATAAGCTCGG 6240
Qy 6241 ACTGCTTAGGCAATGCTCCGCTCGTGGCTAAGGATGTTTGGGATTTGGATATGCA 6300
Db 6241 AGTGTAACCACTCCATGCTCCGTTCTGGCTAAGGACATCTGGGACTGGATATCGAGG 6300
Qy 6301 TGTGTACTCACTTCAAGACTGGCTCCAGTCCAAACTCTCTGCGCGGTTTACCGGAGTCC 6360
Db 6301 TGCTGAGCAGCTTTAAGACTGGCTGAAGAGCCAGCTCATGCCCCAACAATGCTCCGGATTC 6360
Qy 6361 CTTTCTGTATGCCAACCGGGTACAGGGAGTCTGGCGGGGAGCGGCATCATGCAAA 6420
Db 6361 CTTTGTGTCTTGCCAGCGGGGTATAGGGGGTCTTGGCGAGGAGACGGCATTATGCA 6420

Qy 6421 CCACCTGCCATGCGGAGCACAGATCGCCGACATGTCAAAAACGTTTCCATGAGGATCG 6480
Db 6421 CTGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGAGCATGAGGATCG 6480
Qy 6481 TAGGGCTAGAACCTGTGAGCAACACGTGGGACCGAAAGTGTCCCATCAACCGCATACACA 6540
Db 6481 TCGTCTTAGGACCTGTGAGGAACATGTGGAGTGGGAGCTTCCCATTAACCGCTACACCA 6540
Qy 6541 CGGACCTTGCACACCTCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGTGGCTG 6600
Db 6541 CGGGCCCTGTACTTCCCTTCTCGCCGCAACTATAAGTTTCGCGCTGTGGAGGGTGTCTG 6600
Qy 6601 CTGAGGATGACGTGGAGGTTACGCTGTGGGGATTTCCACTAGTTCAGCGGCATGACCA 6660
Db 6601 CAGAGGATACGTGGAGATAAGCGGGTGGGGACTTCCACTACGTATCGGATGACTA 6660
Qy 6661 CTGCAACGTAAGTGCCTATGCGAGTTCGGCCCCCGAATTTCTTCACGAGGTGGATG 6720
Db 6661 CTGCAATCTTAATGCGCGTGCAGATCCATCGCCGAATTTTTCACAGAATTTGAGC 6720
Qy 6721 GAGTGGGTTGCACAGGTACGCTCGCGGTGCAAACTCTTCTACGGGAGGACGTCAGT 6780
Db 6721 GGGTGGCCCTACACAGGTTTGGCCCCCTTGCAGAGCCCTTGTCTCGGGAGAGGTATCAT 6780
Qy 6781 TCCAGGTCGGGCTCAACCAATACTTGGTTCGGGTGCGAGCTCCCATGCGAGCCCAACCGG 6840
Db 6781 TCAGGTAGGACTCAAGGTACCGGTGGGTGCGAATTAACCTTTCGAGGCCGACCGG 6840
Qy 6841 ACGTAAACAGTCTTACTTCCATGCTACCGATCCCTCCCAATTAACAGCAGAGACGGTA 6900
Db 6841 ACGTAGCGGTTCAGCTCCATGCTCATGTATCCCTCCCATATAACAGCAGAGCGCGG 6900
Qy 6901 AGCTAGGCTGGCTAGAGGCTCTCCCTCTTTTACGAGCTCATCAGTACGACGATGT 6960
Db 6901 GGAGAAGGTTGGCGAGAGGGTCAACCCCTTCTATGGCGAGCTCTCTGCGCTAGCAGCTGT 6960
Qy 6961 CTGCGCTTCTTTGAAGCGCATGCTACCTACCACTGACTCCCGGAGCTGACCTCA 7020
Db 6961 CCGTCTCATCTCTCAAGCAACTTGCACCGCAACCATGACTCCCTGACCGGAGTCA 7020
Qy 7021 TCGAGGCCAACCTCTTGTGGCGGAGAGATGGCGGAAACATCACTCGCTGGAGTCA 7080
Db 7021 TAGAGGCTTAACCTCTGTGGAGGAGAGATGGCGGCAACATCACAGGGTTGAGTCAG 7080
Qy 7081 AGAATAGGTAGTAATTTCTGAGCTCTTTGCAACCGCTTCACGCGGAGGGGATGAGAGG 7140
Db 7081 AGAACAAAGTGGTATTTCTGAGCTCTTTCGATTCGCTTGTGGCAGAGAGATGAGCGGG 7140
Qy 7141 AGATATCCGTTCGGCGGAGATCCTGGGAAATCCAGGAAGTTCCCTTCAGCGTTGCCCA 7200
Db 7141 AGGTCTCGTACTCTGCAGAATTTCTGGGAAGTCTCGGAGATTCGCCCGGCGCTGCCG 7200
Qy 7201 TATGGGCAAGCCCGGACTACAATCTCTGCTGTAGAGTCTGGAAGGACCGGACTAG 7260
Db 7201 TCTGGCGCGCGGACTACAAACCCCGCTAGTAGAGACGTGGGAAAAAGCTGACTAGC 7260
Qy 7261 TCCCTCGGTGTACAGGATGCCATTCACCTACCAAGGCTCTTCCAATACCACTC 7320
Db 7261 AACCACTGTGTTCATGGCTGCGGCTACCACTTCCACGCTCCCTCTGTGCTCCGCT 7320
Qy 7321 CACGGAAGAGGAGCGGTTCTGTGACAGAAATCCAAATGTCTTCTGCTTGGCGGAGC 7380
Db 7321 CTCGGAAGAGGAGTACGTTGGTCTCTACCGAATCAACCTATCTACTGCTTGGCGGAGC 7380
Qy 7381 TCGCCACTAAGACCTTCGGTAGCTCCGGATCGTGGCGGTTGATAGCGGCAAGCGGACCG 7440
Db 7381 TTGCCACCAAAAGTTTGGCAGCTCCTCAACTTCGGCATTTACGGGCAACAATACGACA 7440
Qy 7441 CCCTTCGTGACTGGCTCCGACCGGTGACAAAGGATCCGACGTTGAGTTCGTTACTCT 7500
Db 7441 CATCTCTGAGCGCGCCCTTCTGGCTGCCCCCGACTCCGACGTTGAGTTCCTATTCTT 7500

QY	7501	CCATGCCCCCTTTGAAGGGAGCCGGGGAGCCCGGATCTCAGCGACGGGTCTTGGTCTTA	7560
Db	7501	CCATGCCCCCTTTGAAGGGAGCCCTGGGATCCGGATCTCAGCGACGGGTCTAATGTTCGA	7560
QY	7561	CCGT---GAGTGAGGAGGTAGTGAGATGCTGCTGCTCAATGTCTCTATACGTGGA	7617
Db	7561	CGTCAAGTAGTGGGCGGACACCGAGATGCTGCTGCTCAATGTCTTATCTCTGGA	7620
QY	7618	CAGGCGCCCTGTATCAGCCCATCGCTGCGGAGGAAAGTAGCTGCCCATCAACCCGTTGA	7677
Db	7621	CAGGCGCACTCGTCACCCGTCGCTGCGGAGGAAACAAAACTGCCCATCAACGCCACTGA	7680
QY	7678	GCAACTCTTTGCTGCTGCTACCAACAATGCTCTACGCCCAAACTCCCGCAGCGCAAGCC	7737
Db	7681	GCAACTGCTGTACGCCCATCAAACTCTGCTGATTTCCACCACTTCAACGAGTGTGCTG	7740
QY	7738	TCCGGCAGAGAGAGGTCAACCTTTTGACAGATTGCAAGTCTCTGGATGATCAATTACCGGAGC	7797
Db	7741	AAAGGCCAAGAAGTACAAATTTGACAGACTGCAAGTCTCTGGACAGCCATTTACAGGACG	7800
QY	7798	TACTCAAGGAGATGAAGCGGAAGCGCTCCACAGTTAAAGGCTTAAGCTTCTATCTATAGAG	7857
Db	7801	TGCTCAAGGAGGTCAAGCAGCGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGG	7860
QY	7858	AGGCTGCAAGCTGACGCCGCCACATTCGGGCCAAATCCAAATTTGGCTATGGGCCAAAG	7917
Db	7861	AAGCTTGACGCTGACGCCGCCCAATTCAGCCAAATCCAAATTTGGCTATGGGCCAAAG	7920
QY	7918	ACGTCCGNAACCTATCAGCAGAGGCGGTAAACACATCCGCTCGCTGCGGAGGACTTGC	7977
Db	7921	ACGTCCGCTGTCATGCGCAGAAAGCGCGTAGCCCACTCAACTCCGCTGTGAAAGACCTTC	7980
QY	7978	TGGAAGACACTGAAACCACTTACACCACTCATCTGGCAAAAGTGAGGTTTTCTGCG	8037
Db	7981	TGGAAGACAGTGAACCACTAGACACTACCATCATGCGCAAGAACAGGTTTTCTGCG	8040
QY	8038	TCCAACCAGAGAGGGAGCGCAAGCAGCTCGCTTATGCTATTCGACACTGGGAG	8097
Db	8041	TTACGCTGAGAGAGGGGGTCTGAAGCCAGCTCGTCTCTCTGCTGTTCCCGACCTGGCG	8100
QY	8098	TTCTGTATCGGAGAGATGGCCCTTTACGACGTGCTCCACCCTTCCTCAGGCCCTGA	8157
Db	8101	TGCGCGTGTGCGAAGATGGCCCTGTACGACGTGGTTAGCAAGCTCCCCCTGGCCGTGA	8160
QY	8158	TGGCTCTCTATACGGATTTCAATACTCCCGCAAGCAGCGGCTCGAGTTCCTGCTGAATA	8217
Db	8161	TGGAAGCTCTACGATTTCAATPACTCACAGACAGCGGTTGAATTCCTGTCGAAG	8220
QY	8218	CCTGGAATCAAGAAATGCCCCATATGGCTTCTCATATGACACCCGCTGTTTGACTCAA	8277
Db	8221	CGTGGAGTCCAAGAGACCCCGATGGGTTCTCGTATGATACCCGCTGTTTGACTCCA	8280
QY	8278	CGGTCACTGAGAGTGACATGCTGTTGAGGAGTCAATTTACCAATGTTGACTTGGCCC	8337
Db	8281	CAGTCACTGAGAGGACATCGCTACGAGAGGAGCAATTTACCAATGTTGACCTGGACC	8340
QY	8338	CCGAGGCCAGACAGCCATAAGTTCGCTCACAGAGCGCTTTTACATCGGGGCTCCCTGA	8397
Db	8341	CCCAAGCCGGTGGCCATCAAGTCCCTCATGAGAGGCTTTATGTTGGGGCCCTCTTA	8400
QY	8398	CTAACTCAAAAGGGCAGAACTCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGACGA	8457
Db	8401	CCAATTCAGGGGGAAACTCGGCTACCGAGGTGCGCGCAGCGGCTGACTGACAA	8460
QY	8458	CTAGCTCGGTAAATCCCTCAATGTTACTTGAAGGCCACTGACGCTCTGAGGTGCAA	8517
Db	8461	CTAGCTGTGTAAACCCCTCACTTGTCTACATCAAGGCCCGGCGACGCTGTGCGCCGAC	8520
QY	8518	AGCTCCAGACTGCAAGTGTCTGTGAACGAGACGCTTGTCTGTTATCTGTGAAGCG	8577
Db	8521	GGCTCCAGGACTGCACCATGCTCGTGTGCGCAGCACTTAGTCTGTTATCTGTGAAGTG	8580
QY	8578	CGGGAACCCAGAGGATCGGCGGCCCTACGAGCCCTTACGAGGCTATGACTAGGTATT	8637

Db	8581	CGGGGCTCAGGAGAGCGGGGAGCCTGAGAGCCTTACGGAGGCTATGACACGCTACT	8640
QY	8638	CGCCCCCCCCGGGGATCGCCCCCAACCAAGAATACGACCTGGAGCTGATAACAATCATGTT	8697
Db	8641	CGCCCCCCCCGGGGACCCCCCAACCAAGAATACGACTTGGAGCTTATAACATCATGCT	8700
QY	8698	CTTCAATGTGTACGTGCGGACGATGTCATCTGGCAAAAGGGTATACTACTCACCCTG	8757
Db	8701	CTTCAACGTGTGAGTGGCCACGACGGCGCTGGAAGAGGGTCTACTACTTACCCTG	8760
QY	8758	ACCCACCAACCCCTTGCACGGGCTGCGTGGGAGACAGCTAGACACTCCAATCAACT	8817
Db	8761	ACCTACACCCCTTTCGAGAGCGCGTGGGAGACAGACACACTCCAGTCAATT	8820
QY	8818	CTTGGCTAGGCAATATCATATGATGCGCCCACTTATGGGCAAGATGATTTCTGATGA	8877
Db	8821	CTTGGCTAGGCAATATCATATGATGCGCCCACTTATGGGCAAGATGATTTCTGATGA	8880
QY	8878	CTCACTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTGAAGAACCCCTGATTTGTCAGA	8937
Db	8881	CCCATTTCTTTAGCGTCTCATAGCCAGGATCAGCTTGAACAGGCTCTTAACTGTGAGA	8940
QY	8938	TCTACGGGCTTCTACTTCCATTGAGCCACTTGCACCTACCTCAGATCAITGAACGACTCC	8997
Db	8941	TCTACGGGCTTCTACTTCCATTGAGCCACTTGCATCTACTCTCAATCAITCAAAGACTCC	9000
QY	8998	ATGCTCTTAGCGCATTTACACTCCACAGTTACTCTCAGGTGAGATCAATAGGTTGCTT	9057
Db	9001	ATGCTCTTAGCGCATTTACTCTCCACAGTTACTCTCAGGTGAAATCAATAGGTTGCTT	9060
QY	9058	CATGCTCAGAAACTTGGGTTACCACTTGGCAACCTTGGAGACATCGGCGCCGAGATG	9117
Db	9061	CATGCTCAGAAACTTGGGTTACCACTTGGGTTACCACTTGGAGACATCGGCGCCGAGCG	9120
QY	9118	TCCGCTTACGCTTCTCCAGGGGAGGCGGCACTTGTGGCAGATACCTCTTTTA	9177
Db	9121	TCCGCTTACGCTTCTCTCAGAGGAGGAGGCTGCCATATGTGGCAAGTACCTCTTCA	9180
QY	9178	ACTGGGCAAGTAAAGCAAGCTTAAACTCACTCCAACTCCCGGCGCTCCAGCTGACT	9237
Db	9181	ACTGGGCAAGTAAAGCAAGCTTAAACTCACTCCAACTCCCGGCGCTCCAGCTGACT	9240
QY	9238	TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	9297
Db	9241	TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	9300
QY	9298	CCGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	9357
Db	9301	CCGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	9360
QY	9358	TGCTTCCCAACCCGATGAACGGGAGCTTAAACCACTCCAGGC-CTTAAGCCATTTCTGTT	9416
Db	9361	TCCTCCCAACCCGATGAAGGTTGGGTTAAACCACTCCCGGCTCTTAAGCCATTTCTGTT	9420
QY	9417	TT	9476
Db	9421	TT	9480
QY	9477	TT	9536
Db	9481	TT	9540
QY	9537	AGGTCCGTGAGCCGCTGACTGACAGAGAGTGTGATACTTGGCCTCTCTGCAAGATCATGT	9595
Db	9541	AGGTCCGTGAGCCGCTGACTGACAGAGAGTGTGATACTTGGCCTCTCTGCAAGATCATGT	9599

APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 12980 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-09-238-076-5
Query Match 65.3%; Score 6264; DB 3; Length 12980;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 7565; Conservative 0; Mismatches 2030; Indels 53; Gaps 3;
QY 1 GCAGCCCCCTGATGGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAATCTACTG 60
DB 1 GCAGCCCCCTGATGGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAATCTACTG 60
QY 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGCGAGCCCTCCAGGAC 120
DB 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGCGAGCCCTCCAGGAC 120
QY 121 CCCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180
DB 121 CCCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180
QY 181 GACGACCGGTCCTTTCTTTGGATFCAACCCGCTCAATGCTCGAGATTGGGCGTGCCCCC 240
DB 181 GACGACCGGTCCTTTCTTTGGATFCAACCCGCTCAATGCTCGAGATTGGGCGTGCCCCC 240
QY 241 GCGAGACTCTAGCCGAGTAGTGTGGGTGCGGAAGGCGTTGTGTAAGTCTGCTGATAGG 300
DB 241 GCAAGACTCTAGCCGAGTAGTGTGGGTGCGGAAGGCGTTGTGTAAGTCTGCTGATAGG 300
QY 301 GTGCTTGGAGTGCCCCGGGAGGTCTCGTAGACCGGTGACCAATGAGCAGAAATCCCTAAAC 360
DB 301 GTGCTTGGAGTGCCCCGGGAGGTCTCGTAGACCGGTGACCAATGAGCAGAAATCCCTAAAC 360
QY 361 CTCAAGAAAAACCAACCGTAAACCAACCGCGCCCAACGAGACGTCGAAGTTCCCGGGCG 420
DB 361 CTCAAGAAAAACCAACCGTAAACCAACCGCGCCCAACGAGACGTCGAAGTTCCCGGGCG 420

QY 421 GTGCTCAGATCGTTGGTGGAGTTTAACTGTTCGCGCAGGGGCCCCAGGTTGGGTGTCG 480
DB 421 GCGGTCAAGATCGTTGGTGGAGTTTAACTGTTCGCGCAGGGGCCCCAGGTTGGGTGTCG 480
QY 481 GCGGACTAGGAAGGCTTCGAGCGGTGCGAACTCTGTGGAAGGCGAACATATCCCAA 540
DB 481 GCGGACGAGGAAGACTTCGAGCGGTGCGAACTCTGTGGAAGGCGAACATATCCCAA 540
QY 541 AGGCTCGCGACCGAGGCGAGGCGCTGGGCTCAGCGCGGTACCTTTGGGCGCTCTATG 600
DB 541 AGGCAAGTTCGCGCGAGGCGAGGCGCTGGGCTCAGCGCGGTACCTTTGGGCGCTCTATG 600
QY 601 GCAATGAGGCGCTGGGCTGGGCGAGGATGGCTCTCTGTCAACCGCGGCTCCGCGCTAGT 660
DB 601 GCAATGAGGCTGGGCTGGGCGAGGATGGCTCTCTGTCTCCCGGTGGCTCTCGGCTAGCT 660
QY 661 GGGGCCCCACGACCGCGCGGTAGGTGCGTAATCTTTGGGTAAAGTCAATCATACCCCTTA 720
DB 661 GGGGCCCCACAGACCGCGCGGTAGGTGCGTAATCTTTGGGTAAAGTCAATCATACCCCTTA 720
QY 721 CATGGGCTTCGCGGATCTCATGGGTACATTCGCTCGTGGCGCGCCCCCTAGGGGCG 780
DB 721 CGTGGGCTTCGCGGATCTCATGGGTACATTCGCTCGTGGCGCGCCCCCTAGGGGCG 780
QY 781 CTGCCAGGCGCTTCGCGACACGCTGTCCGGGTTCCTGGAGGACGGGTGAATATGCAACAG 840
DB 781 CTGCCAGGCGCTTCGCGACACGCTGTCCGGGTTCCTGGAGGACGGGTGAATATGCAACAG 840
QY 841 GGAATTCGCGGTTGCTTTCTTCTATCTTCCTCTTGGCTGTGCTGTCTGTCTTTGACCA 900
DB 841 GGAATTCGCGGTTGCTTTCTTCTATCTTCCTCTTGGCTGTGCTGTCTGTCTTTGACCA 900
QY 901 TCCAGCTTCGCGTATGAGTGGCGACGCTGTCGGGATATACCATGTGTCAGCAAGACT 960
DB 901 TCCAGCTTCGCGTATGAGTGGCGACGCTGTCGGGATATACCATGTGTCAGCAAGACT 960
QY 961 GCTCAACTCAAGCATTTGTATGAGGCGGAGCTGATCATGATCATCTCCGCGTGGC 1020
DB 961 GCTCAACTCAAGCATTTGTATGAGGCGGAGCTGATCATGATCATCTCCGCGTGGC 1020
QY 1021 TGCCCTGTGTTGAGGAGGTAAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCAGCTCG 1080
DB 1021 TGCCCTGTGTTGAGGAGGTAAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCAGCTCG 1080
QY 1081 CGGCGAGAAATGCCAGGCTCCCACTAGCAATAACGAGCGGACGTCGACTGTCTGCTG 1140
DB 1081 CGGCGAGGAGCGGAACTCCCACTAGCAATAACGAGCTTCGACGCTCATATCGATCTGCTG 1140
QY 1141 GGAAGGCTGCTTTCTGCTCCGCTATGATGAGTGGGATCTCTCGGATCTATTTTCTCG 1200
DB 1141 GGAAGGCTGCTTTCTGCTCCGCTATGATGAGTGGGATCTCTCGGATCTATTTTCTCG 1200
QY 1201 TCTCCAGCTGTTCACTTTCTGCTCCGCGATGAGCAGTGCAGGACTGCAATGCT 1260
DB 1201 TCTCCAGCTGTTCACTTTCTGCTCCGCGATGAGCAGTGCAGGACTGCAATGCT 1260
QY 1261 CAATCTATCCGCGCATGATATGAGTGCACCGCTTGGGATGATGATGAACTGCT 1320
DB 1261 CAATCTATCCGCGCATGATATGAGTGCACCGCTTGGGATGATGATGAACTGCT 1320
QY 1321 CACCTTACAAAGCGCTAGTGTGTCGCAAGTGTCTCCGATCCCAAGCTGTCGTGGACA 1380
DB 1321 CACCTTACAAAGCGCTAGTGTGTCGCAAGTGTCTCCGATCCCAAGCTGTCGTGGACA 1380
QY 1381 TGGTGGCGGGGCCCACTGGGAGTCTCTGGGCGCTTGCCTCACTATTTCCATGGTAGGA 1440
DB 1381 TGGTGGCGGGGCCCACTGGGAGTCTCTGGGCGCTTGCCTCACTATTTCCATGGTAGGA 1440
QY 1441 ACTGGGCTAAGGTTCTGATTTGTTGGGCTACTCTTTGGCGGCTTACGCGGAGACCCACA 1500
DB 1441 ACTGGGCTAAGGTTCTGATTTGTTGGGCTACTCTTTGGCGGCTTACGCGGAGACCCACA 1500
QY 1501 CGACGGGAGGTTGGCGGCGCACACACTCCGCTTCCAGCTCCCTTTCTCATCTCTGGGG 1560

Db 1501 TCACCGGGGAAGTGCAGCGCCGACACACCGCTGGGCTTGTGTCTTACACAGGCG 1560
Qy 1561 CGTCTCAGAAATTCAGCTTGTGAATACCAACGGCAGCTGGACATCAACAGAGACTGCC 1620
Db 1561 CCAAGCAGACATCCAACTGATCAACACCAACGGCAGTTGGACATCAATAGCAGCGCT 1620
Qy 1621 TAAATGCAATGACTCCCTCCAACTGGGTTCTTTTGGCGGCTGTTTACGACACAGT 1680
Db 1621 TGAATGCAATGAAGCCTTAAACCGGCTGGTTAGCAGGCTCTTCTATCAGACAAAT 1680
Qy 1681 TCAACTCGTCCGGTGCAGGAGCGATGCGCAGCTGCGCCCAATGACTGTTGCGCC 1740
Db 1681 TCAACTCTTCAGGCTGCTCAGAGGTTGGCCAGCTGCCAGCGCTTACCGATTTGCCC 1740
Qy 1741 AGGGTGGGCGCCCATCACCTATACCTAAGCTTAACAGCTCGGATCAGAGGCTTATGCT 1800
Db 1741 AGGGTGGGCTCTATCAGTTATGCGCAACGAGAGCGCTCGAGCAACGCGCTACTGCT 1800
Qy 1801 GGCATTACGGCTCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 GGCATTACCGCTCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy 1861 ATTGTTTACCGCCAGCGCTTGTGCTGGGACCAAGGATCGTTCCGCTGCTCCACT 1920
Db 1861 ATTGCTTCACTCCAGCGCCGCTGCTGCTGGGACCAAGGATCGTTCCGCTGCTCCACT 1920
Qy 1921 ATAGTGGGCGGAATGAGACAGCTGATGCTCTTCAACAAACAGCGCTCGGCCACAAG 1980
Db 1921 ACAGTGGGCTGCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Qy 1981 GCAATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 GCAATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Qy 2041 CCCGCTGTAACATCGGGGGGTCGTTAACCGCAGCTTGAATGCTGCCCCCAGGACTGCTTC 2100
Db 2041 CCCCTTGTGTCATCGGAGGGTGGGCAACACACCTTGTCTGCCCCCACTGATTTTTC 2100
Qy 2101 GGAAGCAGCGGCTGCTTACAAATGCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 2101 GCAAGCATCCGGAAGGCACATCTCTCGGTGCGGCTCGGCTGCTGCTGCTGCTGCTGCT 2160
Qy 2161 GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCT 2220
Db 2161 GCATGCTGCACTACCCGTAAGGCTTTGGCACTATCTTGTACCATCAATTTACACCATAT 2220
Qy 2221 TTAAGGTTAGGATGATGTTGGGGGGCTGGAGGACAGGCTCAATGCGGCATGCAATGGA 2280
Db 2221 TCAAAAGTCAGGATGATGCTGGGAGGGTGGAGCAGAGGCTGGAAGCGGCTGCAACTGGA 2280
Qy 2281 CTCGAGGAGCGGCTGTAACCTTGGAGGACAGGATAGGTGAGAACTCAGCGCGCTGCTGC 2340
Db 2281 CGCGGGGGAACGCTGCTGATCTGGAAGACAGGACAGGCTCGAGCTCAGCCATTTGCTGC 2340
Qy 2341 TGTCTACAAAGAGTGGCAGATCTGCCCCCTGCTTTTCAACACCCCTACCGGCTTTATCCA 2400
Db 2341 TGTCCACACACAGTGGCAGGCTCTTCCGTTCTTTTCAACGACCCCTGCGAGCTTTGCTCA 2400
Qy 2401 CTGCTTTGATCCATCTCCATCAGAACTGCTGGAGCTGCAATACCTGTACCGGTGAGGGT 2460
Db 2401 CCGGCTCATCCACCTCCACAGAACTTGTGGAGCTGCAATGCTGCTGCTGCTGCTGCTGCT 2460
Qy 2461 CAGCGCTTGTCTCTTTGCAATCAATGGGAGTACATCTGTTGCTTTTCTCTCTCTCTCT 2520
Db 2461 CAAGCATCGGCTCTCTGGGCAATTAAGTGGAGTACGCTGCTCTCTCTCTCTCTCTCTCT 2520
Qy 2521 CAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Db 2521 CAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Qy 2581 CTTTAGAACTTGTGCTGCTCAATGCGGCTGCTCAATGCGGCTGCTGCTGCTGCTGCTGCT 2640

Db 2581 CTTTGGAGAACTCGTAACTCAATGACGATCCCTGGCCGGAGCGACGGTCTTGTGT 2640
Qy 2641 CTTTCTTGTGTCTTCTCGCGCGCTGTGATCAATTAAGGCGAGGCTGGCTCTCTGGGCGG 2700
Db 2641 CTTTCTTGTGTCTTCTCGCGCGCTGTGATCAATTAAGGCGAGGCTGGCTCTCTGGGCGG 2700
Qy 2701 CGTATGCTTTTATGCGGTATGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Db 2701 TCTACGCTTTCTACGGGATGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Qy 2761 CTTTACGCTTTGACCGGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Db 2761 CATACGCACTGGACACGAGGTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Qy 2821 TATCTTACGCTTGTCCATACCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2880
Db 2821 TGGCGTGTCTCTGTCGCTATTAACAGCGCTACATGCTGCTGCTGCTGCTGCTGCTGCT 2880
Qy 2881 AATACTTTTATCACAGAGCGAGCGCACATGCAAGTGTGGGTCCTCCCTCTCAAGCTTTC 2940
Db 2881 AGTATTTTCTGACAGAGTAGAAGCGAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
Qy 2941 GGGAGGCGCGGATGCT 3000
Db 2941 GGGGGGGCGCGATGCT 3000
Qy 3001 ACATACCAAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
Db 3001 ACATACCAAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
Qy 3061 CGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 3061 TTAAGTTCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Qy 3121 AAGTTCGCGGGGCTCAATGCTCAAAATGCTTCTTCAAGAGCTGGCGCGCTGACAGGTA 3180
Db 3121 AGATGCGCGGAGTCAATGCTGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
Qy 3181 CGTACGTTTATAACCACTTCTTACCCCACTGCGGAGCTGGGCGCACGCGGCGCTAGGAG 3240
Db 3181 CTTATGTTATAACCACTTCTTACCCCACTGCGGAGCTGGGCGCACGCGGCGCTAGGAG 3240
Qy 3241 TTGGGTTGGGCTGAGACCGCTGCTCTCTCTGCGCATGGAGACCAAGCTCATACCTGCG 3300
Db 3241 TGCGCGGCTGTTGGAAACAGTCTCTCTCTGCGAATGGAGACCAAGCTCATACCTGCG 3300
Qy 3301 GAGCAGACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
Db 3301 GGGCAGATACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
Qy 3361 GGAAGAGATATTTTGGGACCGGCTGATGCTGCGAAGGGCAAGGGTGGGAGCTCTCTTG 3420
Db 3361 GCCAGGAGATCTGCTTGGGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Qy 3421 CGCCCATCAGGCTCTCTCCCAACAAACGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db 3421 CGCCCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Qy 3481 TCACAGCGCGGCAAGAAACAGGCTGCAAGGGGAGGTTCAAGTGGTCTTCTACCGCAAC 3540
Db 3481 TGACTGCGCGGCAAGAAACAGGCTGCAAGGGGAGGTTCAAGTGGTCTTCTACCGCAAC 3540
Qy 3541 AATCTTCTGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
Db 3541 AAACTTCTGCGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
Qy 3601 CGAAGACCTTAGCGGCTCAAAAGGCTCAATACCCCAATGCTACCAATGCTAGACCTG 3660
Db 3601 CGAGGACCTTAGCGGCTCAAAAGGCTCAATACCCCAATGCTAGACCTGCTAGACCTG 3660
Qy 3661 ACCTGCTGCTGCGGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3720
Db 3661 ACCTGCTGCTGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3720

Qy	3721	GCTCGGACCTTTACTTGGTCA	GAGACATGCTGATGTCA	TTCCGGTGC	CGCGGAGGCG	3780
Db						
3721		CCTCGGACCTTTACTTGGTCA	GAGGACGCGCATGTCA	TTCCCGTGC	CGCGGAGGCG	3780
Qy	3781	ACAGCGGGGAAGTCTACTCT	CCCCAGGCCCGGTCTCT	CTTAAAGGCTCT	CGGGTG	3840
Db						
3781		ATAGCGGGTAGCTGCTCTTT	CGCCCGGCCCAATTC	TCTATTGAAAGGCTCT	CGGGGG	3840
Qy	3841	GTCCATTGCTTTGCCCTTC	CGGGGACGTCGTGGCGCT	TTCCGGGTGCT	GTGTGCACCC	3900
Db						
3841		GTCCGCTGTTGTGCCCGGG	ACGCCGTGGCCTAITTC	AGGGCCGGGTGT	GCACCC	3900
Qy	3901	GGGGGTTCGGGAAGCGGT	GGACTTATACCCCGTTG	AGTCTATG	GAATACCATCGGT	3960
Db						
3901		GTGAGTGGCTAAGCCGGT	GGACTTATCCCTGTG	GAACCTTAG	AGACAAACATGAT	3960
Qy	3961	CTCCGGTCTTCAAGACAA	CTCAACCCCCCGGTGT	TAACGAGACAT	TCCAAAGTGGCAC	4020
Db						
3961		CCCCGGTGTTCAGGACAA	CTCTCTCCACAGCAGT	GCCACAGAGCTT	CCAGGTGGCCC	4020
Qy	4021	ATCTGCAGCTCTACTTGG	CAGCGGCAAGACCAAGT	GCCGGTGC	GTATGACAGCCC	4080
Db						
4021		ACCTGCAATGCTCCAC	CGGACGGGTAAAGACCA	AGGTCCCGGTGC	GTATGACAGCCC	4080
Qy	4081	AAGGATCAAGTGTCTCGT	CTGAACCCGTCCGTTG	CGGCACCTTAG	GGTTTGGGGCGT	4140
Db						
4081		AGGGCTACAGGTGTTGGT	GTCTCAACCCCTCTGT	TGTGCAAGCTGGGCTT	TGTGCTT	4140
Qy	4141	ATATGTCCAAGGCA	CACGGTATGCACCTAAC	ATCAGAACTGGGGT	TAAGACCAATTACCA	4200
Db						
4141		ACATGTCCAAGGCCAT	GGGTTGATCCTAATAT	CAGGACCGGGT	GTAGAACAATTACCA	4200
Qy	4201	CGGGGGCTCCATTACGT	ACTCCACCTATGGCA	AGTTCCTTGCCG	AGGTGGCTGTCTG	4260
Db						
4201		CTGGCAGCCCATACG	TACTCCACCTACGGCA	AGTTCCTTGCCG	AGGTGGCTGTCTG	4260
Qy	4261	GGGGCGCTATGACAT	CATAATATGTGATG	ATGCCACTCAACTG	ACTCCGACTACCATCT	4320
Db						
4261		GAGTGTCTTATGACAT	ATAATTTGTGAGAGT	GCCACTCA	CGGATGCCACATCCATCT	4320
Qy	4321	TGGGCATCGGCACAGT	CTTGACCAAGCGGAG	AGCGCTGGAGCGCGCT	CGTCTGTCTG	4380
Db						
4321		TGGGCATCGGCACAGT	CTCTTGACCAAGCAG	AGACTCGGGGGCGAG	ACTGTTGTGCTG	4380
Qy	4381	CCACGGCTACCTTCGG	GGATCGGTTACCGTG	CCACACCCCAATAT	TCGAGGAATAGGCC	4440
Db						
4381		CCACTGTCTACCCCT	CCGGGCTCCGTCACT	GTGTGCCATCTCTAA	CATTCGAGGAGGTTGCTC	4440
Qy	4441	TGTTCCAAATGGAGAG	ATCCCTTCTATGCA	AGGCCATCCCATTT	AGCCATTAAGGCCATCAAGG	4500
Db						
4441		TGTTCCACCGAGAG	ATCCCTTTTACGCA	AGGCTATCCCCCT	CGAGGTATCAAGG	4500
Qy	4501	GGGGAGGCATCTCAT	TTTTCGCCATTTCCA	AGAGAAATGTG	ACGAGCTCGCCGCAAGC	4560
Db						
4501		GGGGAAGACATCTCAT	CTTCTGCCACTCA	AGAGAAAGTGG	ACGAGCTCGCCGCAAGC	4560
Qy	4561	TGACAGGGCTCGGACT	GAAACGCTGTAGCA	ATATACGGGGCTT	GTATGTTCGTCATAC	4620
Db						
4561		TGTTGCGCATTTGGCA	TCAATCAATCCCGT	GGCTACTACTACCG	GGTCTTTGACGTGTCTGTCATCC	4620
Qy	4621	CGCCTATCGGAGAG	CTGCTGTGTGGCA	CAGACGCTCTAAT	CACGGTTTCCACCGCG	4680
Db						
4621		CGACAGCGGGGATG	TGTGCTGTGTGC	ACCGATGCTCTCAT	GACTTGGCTTTACCGGG	4680
Qy	4681	ATTTTGACTCAGTGA	TGACTGCAATACAT	GTGTCAACCG	TTTTCACCGGCTT	4740
Db						
4681		ACTTCGACTCTGTG	ATGACTGCAACAGT	GTGTCACTCAG	ACAGTGCATTCAGCTTG	4740
Qy	4741	ATCCCACTTCCATTTG	AGACGACCGGTGCC	CCAAAGCGGGGT	GTGCGGCTTCGCAAC	4800
Db						
4741		ACCTTACCTTTACCA	TGTGACCAACAG	CGCTCCGCCAGG	ATGCTGTCTCTCAGGACTCTCAAC	4800

Qy	4801	GGCGAGGTGAAGACTGCGACGGGTAGAGTGGGATCTACAGGTTTGTGATCTCCAGGAGAAC	4860
Db	4801	GCCGGGCAGGACTGCGACGGGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGGGGAGC	4860
Qy	4861	GGCCCTCGGGCATGTTCGAATCTTTCGGTCCCTGTGTGAGTGCTATGACGCGGGCTGTGCTT	4920
Db	4861	GCCCTCTCCGCATGTTTCGACTCGTCCGTCTCTGTGAGTGCTATGACGCGGGCTGTGCTT	4920
Qy	4921	GGTATGAGTCTACGCCCGCTGAGACCTTCGGTTTAGTGTGGGGCTTACTCTAAATACACCAG	4980
Db	4921	GGTATGAGTCTACGCCCGCGGAGACTACAGTTAGGCTACGAGGCGTACATGAACACCCCGG	4980
Qy	4981	GGTTGCCGCTCTGCCAGGACCAATCTCGAGTTCTGGGAGAGCGTCTTACAGGCGCTCACCC	5040
Db	4981	GGCTTCCCGTGTGCCAGGACCAATCTTGAATTTTGGGAGGGCGTCTTTATCGGGCTCACTC	5040
Qy	5041	ACATAGATGCCACTTCTGTCTCCAGACTTAACAGGACGAGAGCAACTTTCCTTACTCTGG	5100
Db	5041	ATATAGATGCCACTTCTTATCCACACAAAGCAGAGTGGGGAGAATTTCTCTTACTCTGG	5100
Qy	5101	TGCGCATATCAAGCTACAGTGTGGCGCAGGCGTCAAGCTCCACCTCCATCGTGGGACCAAA	5160
Db	5101	TAGCGTACCAAGCCACCGTGTGGCGTATAGGCTCAAGCCCTCCCCCATCTGGGACGAGA	5160
Qy	5161	TGTGGAAGTCTCTCATACGGCTGAACCTACACTGCACGGGCCAAACACCCCTCTCTGTATA	5220
Db	5161	TGTGGAAGTGTTTGATCCCGCTTTAAACCCACCCCTCCATGGGCCAAACACCCCTGCTATACA	5220
Qy	5221	GGCTAGAGCGGTCCAAAATGAGGTCATCTCTACACACCCCATAACTAAATATCATCATGG	5280
Db	5221	GACTGGCGCGTGTTCAGAAATGAAGTCAACCTGACCGCAACCAATCACCAATATCATCATGA	5280
Qy	5281	CATGCATCTCGGCTGACCTGGAGGTCGTACTAGCACCTTGGTGCTGGTAGGGGAGTCC	5340
Db	5281	CATGCATCTCGGCCGACCTGGAGGTCGTACGAGCACCTGGGTGCTCGTTGGCGGCGTCC	5340
Qy	5341	TTGCAGCTTTGGCCGCATACTGCTCTACGACGACGAGTGTGGTCATTGTGGGCGAGATCA	5400
Db	5341	TGGCTGCTCTGGCCGCGTATTGCTGTCAACAGGCTGCGTGGTCATAGTGGCGAGGATTG	5400
Qy	5401	TCATTGTCGGGGAAGCAGCTGTGTTCTCCGACAGGGAAGTCTCTTACCAGAGTTTCGATG	5460
Db	5401	TCTTTGTCGGGGAAGCCGCCAATTTATACCTGACAGGGAGGTTCTCTACAGAGGTTTCGATG	5460
Qy	5461	AGATGGAAGTGTGCTCTCAACAATCTCTTACATCGAGCAGGGAATCGACGTGCGCCGAGC	5520
Db	5461	AGATGGAAGAGTGTCTCTCAGCATTAACGTTACATCGAGCAAGGATGATGCTCTGCTGAGC	5520
Qy	5521	AAATTCAAGCAAAAGCGGCTCGGTTGTGCAAAACGGCCACCAAGCAAGCAGGAGGCTGCTG	5580
Db	5521	AGTTCAAGCAGAAGGCCCTCGGCTCTCTGAGACCCGCTCCCGCAAGCAGAGGTTTATCA	5580
Qy	5581	CTCCGCTGGTGGAGTCCAAAGTGGGAGCCCTTGAGACCTTCTGGGCGAAGCACAATGTGGA	5640
Db	5581	CCCTGCTGTCTCAGACCAACTGGCGAGAACTCGAGGTCTTCTGGGCGAAGCACAATGTGGA	5640
Qy	5641	ATTTTCATCGGGAATACGTACTAGCAGCTTATCCACTCTGCTCTGGAACCCCGCGA	5700
Db	5641	ATTTTCATCGTGGGATACAATCTTGGCGGGCTGTCAAGCTGCTCTGGTAAACCCCGCA	5700
Qy	5701	TAGCATCATTTGATGGCATTTTACAGTCTTATCACTAGCCCCGCTCACCAACCAAAACCC	5760
Db	5701	TTGCTTCATTTGATGGCTTTTACAGCTGCCGTACACAGCCCACTAAACCACTGGCCAAACCC	5760
Qy	5761	TCCTGTTTAACTTTGGGGGATGGGTGGCTGCCAACTCGCTCTCTCCAGCGCTCGGT	5820
Db	5761	TCCTCTTCAACATATTGGGGGGTGGGTGGCTGCCAGCTCGCCGCCCGCCCGTCCCGCTA	5820
Qy	5821	CAGCTTTCTGGGCGCGCATCGCCGAGCGGCTGTGTCAGCATAGGCTCTGGGAAGG	5880
Db	5821	CCGCTTTGTGGGCGCTGGCTTAGCTGGCGCCGCTTACGTCGGCAGGCTTGGATCTGGGGAAG	5880
Qy	5881	TGCTCTGTGGAATCTCTTGGCGGGCTATGGGCGAGGGTAGCCGGCGCACTCGTGGCTTTA	5940

5881	Db	 TCTCTGTGGACATCTTGTGAGGGTATGGCGCGGCGTGGCGGAGCTCTTGTAGCGTTCA	5940
5941	Qy	 AGGTCATGAGCGCGAGGTGCCCTCCACCGAGGACCTGGTCAACTTACTCCCTCGCATCC	6000
5941	Db	 AGATCATGAGCGGTGAGGTCCCTCCACGAGGAGACCTGGTCAATCTGCTGCCCGCCATCC	6000
6001	Qy	 TCTCTCGTGTGCCCTGTGTCTCGGGTCTGTGTGCGCAGCAATPACTCTGCGTCGCGACGTGG	6060
6001	Db	 TCTCGCTTGGAGCCCTTGTAGTCGGTGTGTCTGCGCAGCAATPACTCTGCGCGCACGTTG	6060
6061	Qy	 GCCCGGAGAGGGGGCTGTGCACTGTGATGAAACCGGCTGATAGGTTTCGCTTCGGGGGTA	6120
6061	Db	 GCCCGGCGAGGGGGCAGTGCATATGGATGAACCGGCTAATAGCTTCGCTCCCGGGGA	6120
6121	Qy	 ACCACTCTCCCTACGCACTATGTCCTGAGAGCGACGCTGCGACGACGTGTCACTCAGA	6180
6121	Db	 ACCATGTTTCCCCACGCACTAGCTGCGGAGAGCGATGACGCCGCCGCGTCACTGCCA	6180
6181	Qy	 TCTCTCTAGCCCTTACCATCACTCAACTGTGAAGCGGTCCACCACTGGATTAATGAGG	6240
6181	Db	 TACTCAGCAGCCTCATCTGTAAACCCAGCTCTGAGGCGACTGCATCAGTGGATTAAGCTCGG	6240
6241	Qy	 ACTGCTCTACGCCCATGCTCCGGTCTGFGTAAAGGATGTTTGGGATTTGGAATGACACGG	6300
6241	Db	 AGTGTACCACTCCATGCTCCGGTCTCTGGCTAAGGGACATCTGGGACTGGATATGGGAGG	6300
6301	Qy	 TGTGTACTGACTTCAAGACCTGCTCCAGTCCAAATCTCTGCGCGGTTTACCGGGAGTCC	6360
6301	Db	 TGTGTAGCGACTTTAAGACCTGCTGAAGCCAAAGCTCATGCCACAATCGCTCTGGGATTC	6360
6361	Qy	 CTTTTCTGTCTATGCAACCGGGGTACAAGGAGTCTTGGCGGGGGGACGGCATATGCAAA	6420
6361	Db	 CCTTTGTCTCTGCCAGCGGGGTATAGGGGGGTCTGGCGAGGAGACGGCAITATGCACA	6420
6421	Qy	 CCAACCTGCCCATCGGGAGCAGATGCCGGACATGTCAAAAACGGTTTCCATGAGGATCG	6480
6421	Db	 CTCGCTGCCACTGTGGAGCTGAGATCACTGGAATGTCAAAAACGGGACGATGAGGATCG	6480
6481	Qy	 TAGGGCTAGAACCTGACGCAACACGTGCGCAGCGAGTTCGCCCATCAACGCAATACACCA	6540
6481	Db	 TCGGTCTTAGGACCTGCAGGAACATGTGGAAGTGGGAGCTTCCCATTTAAGCCTTACACCA	6540
6541	Qy	 CGGGACCTTGCACACCTCCCGCGCGCCAACTATTCCAGGGGCTATGCGGGTGGCTG	6600
6541	Db	 CGGGCCCTGTACTCCCTTCTCTGCGCGGAACCTATAAGTTTCGGCTGTGAGGGTGTCTG	6600
6601	Qy	 CTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGCAATGACCA	6660
6601	Db	 CAGAGGAATACGTGGAGATAAGCGGGTGGGGACTTCCACTACGTATCGGTATGACTA	6660
6661	Qy	 CTGACAAAGTTAAAGTGCCTATGCCAGTTCGGGCCCCCGAATCTTTCACGGAGGTGGATG	6720
6661	Db	 CTGACAATCTTAAATGCCCCGTGCCAGATCCCATCGCCCCGAATTTTTCACAGAAATTGGACG	6720
6721	Qy	 GAGTGCGGTTGCACAGGTACGCTCCGGCGTGCAAACTCTTTCTACGGGAGGAGCTCACGT	6780
6721	Db	 GGGTGCGCCTACATAGTTTTCGGCCCCCTTGCAGACCTTGTCTCGGGAGGAGGTATCAT	6780
6781	Qy	 TCCAGGTCCGGCTCAACCAATPACTTGTGTGGGTGCGAGCTCCCATGCGAGCCCCGACCGG	6840
6781	Db	 TCAGAGTAGGACTCCACGAGTACCGGTGGGTGCGCAATTACCTTTCGAGCCCCGAAACCGG	6840
6841	Qy	 ACGTAAACAGTGCCTTACTTCATGTCTACCGATCCCTGCCACATTTACAGCAGACAGCGCTA	6900
6841	Db	 ACGTAGCCGTGTGACGTCCATGTCTCACTGATCCCTCCCATATATACAGCAGAGCGCCG	6900
6901	Qy	 AGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTTAGCCAGCTCATCAGCTAGCCAGTTGT	6960
6901	Db	 GGAGAAGGTTGGCAGAGGGGTCACCCCCTTCTATGGCCAGCTCTCTGGCCAGCCAGCTGT	6960
6961	Qy	 CTGGCCCTTCTTTGAAGCGCAGATGCACTACCCACCATGACTCTCCCGGACGCTCACTCA	7020

6961	CGGCTCCATCTCTCAAGGCAACTTTGCACCGCCAAACCATGACTCCCTCTGACGCCGAGCTCA	7021	TCGAGGCCAACCTCTTTGTGGCGGCGAGAGATGGCGGAAACATCACTCGCGTGGAGTCAG
7021		7021	
7021	TAGAGGCTAACCTCTCTGTGGAGCGCAGGAGATGGCGCGCAACATCACACAGGTTTGAGTCAG	7081	AGAAATAGGTAGTAATTTCTGGACTCTTTTCGAACCGCTTTCACGCGGAGGGGGATGAGAGGG
7081		7081	
7081	AGAAACAAGTGGTGATTTCTGGACTCTCTTGATCCGCTTTGTGGCAGAGGAGGATGAGCGGG	7141	AGATATTCGTCGCGCGGCGAGATCCTTCGGAATAATCCAGGAAGTTTCCCTCAGCCGTTCGCCCA
7141		7141	
7141	AGTCTTCGCTACCCGCAGAAATTTCTGGGAACTCTCGAGATTCGCGCGGCCCTGCCCG	7201	TATGGCAGCCCGGACTACAATTCCTCACTGCTAGAGTCCTGGAAGGACCCGGACTACG
7201		7201	
7201	TTTGGCGGCGCGGACTACAACCCCGCTAGTAGAGACGTGGAATAAACCTCTGACTACG	7261	TCCTTCGGTGTACAGGATGCCATTTGCCACTACCAAGGCTCTCTCCAATACCACTTC
7261		7261	
7261	AACCACTGTGTGTCCATGGCTGCGCGCTACCACTCTCACGGTCCCTCTGTGCTCCCTCCG	7321	CAGCGAAGAGAGCGTTGTCTCAGCAAAATCCAAATGTCTCTTGCTCTTCGCGGAGC
7321		7321	
7321	CTCGGAAAAAGCGTAGCGTGGTCTCTCACGAAATCAACCTATCTACTGCTCTTGGCGAGC	7381	TCGCCACTAAGACCTTTCGGTAGCTCCGGATTCGTGCGCGTTGTATAGCGGCACCGCGCACG
7381		7381	
7381	TTGCCACCAAAAGTTTGGCAGCTCTCTCAACTTTCGSCATTTACGGGCGACAATACGACAA	7441	CCCTTCCTGACTGGGCTTCGACGACGCGTGACAAAGGATCCGAGCTTGAGCTGCTACTCTCT
7441		7441	
7441	CATCTCTGAGCGCGCGCCCTTCTTGGCTGCGCCCGGACTCCGACGTTTGAGTCTCTATTCTT	7501	CCATGCCCCCTTTGAAGGGGAGCGCGGGACCCCGATCTCAGCGACGGGTCTTGCTCTA
7501		7501	
7501	CCATGCCCCCTTGGAGGGGAGCGCTGGGATTCGGATCTCAGGACCGGTCATGCTCGA	7561	CGGTCACTAGTGGGCGGACACGGAAGATGTCGTGTGCTGCTCAAATGCTTATTCTCTGGA
7561		7561	
7561	CGGTCACTAGTGGGCGGACACGGAAGATGTCGTGTGCTGCTCAAATGCTTATTCTCTGGA	7617	CGGT - - - GAGTGAGGAGGCTAGTGAGGATGTCGTCTGCTCTCAATGTCCTATACGTGGA
7617		7617	
7617	CGGTCACTAGTGGGCGGACACGGAAGATGTCGTGTGCTGCTCAAATGCTTATTCTCTGGA	7677	CAGCGCCCTTGATCAGGCCATGGCTTCGGAGGAAAGTAAGCTTGCCCATCAACCCGTTGA
7677		7677	
7677	CAGCGCACTCGTCACCCCGTGGCGTCGGGAAGAACAAAAAATCTGCCATCAACGCACTGA	7690	GCAACTTTTGTCTGCTCACCACAACATGCTTACGCGCAACAAATCCGAGCGCAAGCC
7690		7690	
7690	GCAACTTTTGTCTGCTCACCACAACATGCTTACGCGCAACAAATCCGAGCGCAAGCC	7737	GCAACTTTTGTCTGCTCACCACAACATGCTTACGCGCAACAAATCCGAGCGCAAGCC
7737		7737	
7737	CGCGCAGAGAAGGTCACCTTTTGACAGATTGCAAGTCTGGAATGATCATTTACCGGAGC	7797	TACTCAAGAGAGATGAAGGGGAAAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAG
7797		7797	
7797	TACTCAAGAGAGATGAAGGGGAAAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAG	7857	TGCTCAAGAGGTCAAAGCAGCGCGTCAAAAGTGAAGGCTTAATCTGCTATCCGTAGAGG
7857		7857	
7857	TGCTCAAGAGGTCAAAGCAGCGCGTCAAAAGTGAAGGCTTAATCTGCTATCCGTAGAGG	7917	AGGCTCGCAGCTGACGCCCCCATTTTCGCGCAATCCAAATTTGGCTATGGGGCAAGG
7917		7917	
7917	AGGCTCGCAGCTGACGCCCCCATTTTCGCGCAATCCAAATTTGGCTATGGGGCAAGG	7977	ACGTCCCGAACCTATCAGCAGGGCCGTTAACACATCCGCTCCGCTGTGGGAGGACTTGC
7977		7977	
7977	ACGTCCCGAACCTATCAGCAGGGCCGTTAACACATCCGCTCCGCTGTGGGAGGACTTGC	8037	TGGAAGACAGTGTAACACCAATAGACACTACCATCATGGCCAAAAGTGAGGTTTTCTGCG
8037		8037	
8037	TGGAAGACAGTGTAACACCAATAGACACTACCATCATGGCCAAAAGTGAGGTTTTCTGCG	8097	TCCAACAGAGAAAGGGAGCGCAAGCCAGCTTCGCTTATCGTATTTCCACACCTGGGAG
8097		8097	
8097	TCCAACAGAGAAAGGGAGCGCAAGCCAGCTTCGCTTATCGTATTTCCACACCTGGGAG	8100	TTACGCTTGAAGAGGGGGGTGTAAGGCAGCTCTGCTCTCATCTGTTTCTCCGACCTGGGCG
8100		8100	
8100	TTACGCTTGAAGAGGGGGGTGTAAGGCAGCTCTGCTCTCATCTGTTTCTCCGACCTGGGCG		

QY 8098 TTCTGTATGCGAGAGATGGCCCTTTTACGAGCTGGTCTTCAACCCCTTCTCCTCAGGCGCTGA 8157
Db 8101 TCGCGCTGTGCGAGAGATGGCCCTGTACGAGCTGGTTAGCAAGCTCCCCCTGGCCGCTGA 8160
QY 8158 TGGGCTCTCTATACGAGATTTCAATACTCCCCAAGCAGCGGTGAGTTCTCTGTGTAATA 8217
Db 8161 TGGGAAGCTCTTACGAGATTTCAATACTCACCAGGACAGCGGTTGAATTTCTCTGTGCAAG 8220
QY 8218 CTTGGAATCAAGAAATGCCCCCTATGGGCTTCTCATATGACACCCGCTGTTTGTGACTCAA 8277
Db 8221 CTTGGAAGTCCAAGAAAGACCCCGATGGGTTCTCGTATGATACCCGCTGTTTGTGACTCCA 8280
QY 8278 CGTCACTGAGAGTACATCTGCTGTGAGGAGTCAATTTACCAATGTTGACTTGGGCC 8337
Db 8281 CAGTCACTGAGAGCACATCCGTACGAGAGGCAATTTTACCAATGTTGACTTGGACC 8340
QY 8338 CCGAGGCCAGACAGGCCATTAAGTCTGCTCACAGAGCGGCTTTTACATCGGGGTCGCCCTGA 8397
Db 8341 CCCAAGCCCGGTGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTGGGGCCCTCTTA 8400
QY 8398 CTAACCTAAAAGGCGAGAACTCGCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGACGA 8457
Db 8401 CCAATTTCAAGGGGGGAAATCTCGGCTACCGCAGGTGCGCGCAGCGCGGTACTTGACAA 8460
QY 8458 CTAGCTGGGTTAATACCTCTCACATGTTACTTTGAAGGCCACTGCGAGCTGTGAGCTGCAA 8517
Db 8461 CTAGCTGTGTTAACAACCTCTCACTGCTACATCAAGGCCCGGCGAGCTGTGAGCGCGAG 8520
QY 8518 AGCTCCAGGACTGACGATGCTCGTGAACGGAGAGACGCTTCTGCTTATCTGTGAAGCG 8577
Db 8521 GSCCTCCAGGACTGCACCATGCTCGTGTGTTGGGACGACTTACTGCTTATCTGTGAAGTG 8580
QY 8578 CGGGAACCCAGAGAGATGCGCGGCCCTTACGAGCCTTACGAGGCTATGACTAGGTATT 8637
Db 8581 CGGGGCTTCAGGAGGACCGCGGAGCCTGAGAGCCTTACGAGGCTATGACAGGTTACT 8640
QY 8638 CGCCCCCCCCGGGATCGCCCCAACAGATACGACTGAGCTGATACATCATGTT 8697
Db 8641 CGCCCCCCCCGGGACCCCCAACACAGATACGACTTGGAGCTTAAACATCATGCT 8700
QY 8698 CTTCAATGTGTGCTGCGGACGATGCTATCTGGCAAGAGGTATACTACCTCACCCCGTG 8757
Db 8701 CTTCAACGTGTGCTGCGGACGAGCGGCTGGAAGAGGTCTACTACCTTACCCGCTG 8760
QY 8758 ACCCACACCCCTTGTGACGGGTGCTGGGAGACAGCTAGACACACTCCAACTCAACT 8817
Db 8761 ACCCTACAACCCCTTGTGAGAGCGCGTGGGAGACAGACACACTCCAGTCAATT 8820
QY 8818 CTTGGCTAGGCAATATCATGATGATGCGCCCAACCTATGGCAAGGATGATTCTGATGA 8877
Db 8821 CTTGGCTAGGCAACATAATCATGTTTGGCCCCACACTGTTGGCGAGGATGATGATGA 8880
QY 8878 CTCACCTTTTCTCCATCTCTAGCTCAAGAGCAACTTGAAGAGCCCTGGATTGTGAGA 8937
Db 8881 CCCATTTCTTAGCTGCTCTATAGCCAGGATCAGCTTGAACAGGCTCTTAAGTGTGAGA 8940
QY 8938 TCTAGCGGCTGTGCTACTCCATGAGCCACTTGAACCTACCTCAGATCATGTAACGACTCC 8997
Db 8941 TCTAGCGGCTGTGCTACTCCATAGAACCACTGGATCTACCTCAATCATTTCAAGACTCC 9000
QY 8998 ATGGCTTAGCGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGTGGCTT 9057
Db 9001 ATGGCTTAGCGCAATTTTCACTCCACAGTTACTCTCCAGGTGAAATCAATAGGTGGCG 9060
QY 9058 CATGCTCAGGAAACTTGGGGTACCACTCTGCAACCTTGGAGACATCGGGCCAGAAAGTG 9117
Db 9061 CATGCTCAGAAACTTGGGGTCCCGCTTGGAGCTTGGAGACACCGGGCCCGGAGCG 9120
QY 9118 TCCGCGCTAAGCTACTGTCCAGGGGGGAGGGCCGCTCTTGTGCGAGATACCTCTTTTA 9177
Db 9121 TCCGCGCTAGGCTTCTGTCCAGAGGAGGCGGGCTGCCATATGTGCAAGTACTCTTCA 9180

QY 9178 ACTGGCAGTAAGACCAAGCTTAAACTCACTCAATCCCGCGCGCTCCAGCTGGACT 9237
Db 9181 ACTGGCAGTAAGAACAAAGCTCAAACTCACTCAATAGCGCGCTGGCGGCTGGACT 9240
QY 9238 TGTCTGGCTGTGCTGCTGCTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTG 9297
Db 9241 TGTCCGGTTGGTTTACGCGCTGGCTACAGCGGGGAGACATTTATCACAGCGTGTCTCATG 9300
QY 9298 CCCGACCCCGCTGCTTCCGTTGTGCTTACTTCTTCTAGGGGTAGGCATTTACC 9357
Db 9301 CCCGCCCCCGCTGCTTCTGTTTGGCTTACTTCTGCTGCGAGGGTAGGCATCTACC 9360
QY 9358 TGCTCCCCCAACCGATGAACGGGAGCTAAACCACTCCAGC-CTTAAAGCAATTTCTGTTT 9416
Db 9361 TCCTCCCCAACCGATGAAGTTGGGTAAACACTCCGCGCTCTTAGGCCAATTTCTGTTT 9420
QY 9417 TT 9476
Db 9421 TT 9480
QY 9477 TTTCTTTT-----TCC 9487
Db 9481 TTTTTTTTCTTCTTTTCTTTTCTTTT 9540
QY 9488 CTCTTTAATGCTGGCTCCATCTTAGCCCTAGTCAAGCTAGCTGAAAGGTCCGCTGAG 9547
Db 9541 CTCTTTAATGCTGGCTCCATCTTAGCCCTAGTCAAGCTAGCTGAAAGGTCCGCTGAG 9600
QY 9548 CCGCATGACTCAGAGAGTCTGATGACTGGCTCTCTGCAGATCATGT 9595
Db 9601 CCGCATGACTCAGAGAGTCTGATGACTGGCTCTCTGCAGATCATGT 9648

RESULT 10

US-09-995-937-5
; Sequence 5, Application US/0995937
; Publication No. US20030028010A1
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; COUNTRY: USA
; STATE: MO
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,937
; FILING DATE: 28-Nov-2003
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,756
; FILING DATE: 04-May-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

		TOPOLOGY: linear		Query Match		65.3%; Score 6264; DB 3; Length 12980;			
		MOLECULE TYPE: cDNA		Best Local Similarity		78.4%; Pred. No. 0;			
		HYPOTHETICAL: NO		Matches 7565; Conservative		0; Mismatches 2030; Indels 53; Gaps 3;			
		ANTI-SENSE: NO							
		SEQUENCE DESCRIPTION: SEQ ID NO: 5:							
		US-09-995-937-5							
QY	1	GCACGCCCTTGATGGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAACCTACTG	60	Db	901	TGCCCCCTTCAGCCTACCAAGTGGCAATTCTCTGGGGCTTTTACCATTGTCAACAATGATT	960		
Db	1	GCACGCCCTTGATGGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAACCTACTG	60	QY	961	GCTCCAACCTAAGCAATGTGTATGAGGAGCGGAGCGTGTATCATGTATCATCTCCGGGTGCG	1020		
QY	61	TCTTACGCAAGAAAGCGTTAGCCATGGCGTTAGTATGATGATGCTGCGAGCCTCCAGGAC	120	Db	961	GCCTAACTCGAGTATTGTGTACGAGGCGGCGGATGCCATCTCTGCACACTCCGGGGTGTG	1020		
Db	61	TCTTACGCAAGAAAGCGTTAGCCATGGCGTTAGTATGATGATGCTGCGAGCCTCCAGGAC	120	QY	1021	TGCCCCCTGTTCAGGAGGGTAACAGCTCCCGTTGCTGGGTAGCGTCACTCCCAAGCTCG	1080		
QY	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG	180	Db	1021	TCCCTTGGCTTCGAGGGTAACCGCTCGAGGTGTGGGTGGCGGTGACCCCCACCGTGG	1080		
Db	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG	180	QY	1081	CGGCCAGGAATCCAGCGTCCCACCTACGACAATACGAGCCACGCTCGACTTCTCTCTG	1140		
QY	181	GACGACGGGTCTTCTTGTGATCAACCGCTCAATGCTCGGAGATTGGGCGTGCCTCC	240	Db	1081	CCACAGGAGACGGCAAACTCCCAACACGAGCTTCGAGTATCATATCGATCTGCTTGTG	1140		
Db	181	GACGACGGGTCTTCTTGTGATCAACCGCTCAATGCTCGGAGATTGGGCGTGCCTCC	240	QY	1141	GGACGCTCTCTTCTGCTCCGCTATGTACGTGGGGATCTCTGCGGATCTATTTTCTCTG	1200		
QY	241	GCAGACTCTAGCCGAGTAGTGTGGTGGCGAAGGCCCTTGTGTACTTGCCTGATAGG	300	Db	1141	GGAGCGCCACCTCTGCTCGGCCCTCTACGTGGGGACCTGTGCGGGTCTGTCTTCTTTC	1200		
Db	241	GCAAGACTCTAGCCGAGTAGTGTGGTGGCGAAGGCCCTTGTGTACTTGCCTGATAGG	300	QY	1201	TCTCCAGCTGTTTACCTTCTCTCCAGGCGCCACTGCGACGACGCAAGACTGCAATTTGTT	1260		
QY	301	GTCTTTCGAGTCCCGGAGGTCTCGTAGACCGTGCAATGAGACGAACTCTTAAC	360	Db	1201	TTGGTCAACTGTTTACCTTCTCTCCAGGCGCCACTGCGACGACGCAAGACTGCAATTTGTT	1260		
Db	301	GTCTTTCGAGTCCCGGAGGTCTCGTAGACCGTGCAATGAGACGAACTCTTAAC	360	QY	1261	CAATCTATCCGGCCATGTATCAGGTCAACGATGCTTGGGATATGATGATGAATGAACTGTT	1320		
QY	361	CTCAAGAAAAACCAACGTTAACCAACCGTCCGCCACGAGCGTCAAGTTCCCGGGTG	420	Db	1261	CTATCTATCCGGCCATATAACGGGTCTATCGCATGGCATGGGATATGATGATGAATGTT	1320		
Db	361	CTCAAGAAAAACCAACGTTAACCAACCGTCCGCCACGAGCGTCAAGTTCCCGGGTG	420	QY	1321	CACCTACACAGCCCTAGTGTGCTGCGAGTGTCTCCGATCCCAAGCTGTGCTGAGACA	1380		
QY	421	GTGTCAGATCGTGTGGAGTTTACCTGTTCGCGCAGGGGCCCGAGGTGGGTGTGC	480	Db	1321	CCCTACGGCAGCGTTGGTGTAGCTCAGTCTCGGATCCCAAGCCATCATGAGACA	1380		
Db	421	GTGTCAGATCGTGTGGAGTTTACCTGTTCGCGCAGGGGCCCGAGGTGGGTGTGC	480	QY	1381	TGCTGGCGGGGCCACTGGGAGTCTTGGCGGCGCTTGCCTACTATTTCCATGTTAGGA	1440		
QY	481	GCAGACTTAGAAGGCTTCCGAGCGTTCGAACTCGTGAAGGCGACAACTATCCCA	540	Db	1381	TGATCGTGTGTCTACTGGGAGTCTTGGCGGCATAGCGTATTCTTCATGTTGGGA	1440		
Db	481	GCAGACTTAGAAGGCTTCCGAGCGTTCGAACTCGTGAAGGCGACAACTATCCCA	540	QY	1441	ACTGGCTAAGTGTCTGATTGTGGCGCTACTCTTTGCGGCGGTGTAGCGGGAGACCCACA	1500		
QY	541	AGGCTCGCGCACCGGAGGCGAGGCTTGGGCTACGCGCGGTAACCTTTGGCCCTCTATG	600	Db	1441	ACTGGGCGAAGTCTTGGTGTCTGCTGCTGCTATTGTCGGCGCTCGACGCGGAAACCCACG	1500		
Db	541	AGGCTCGCGCACCGGAGGCGAGGCTTGGGCTACGCGCGGTAACCTTTGGCCCTCTATG	600	QY	1501	CGACGGGAGGGTGGCGGCCACACCACTCCGCGGTTCAGTCCCTTTCTCATCTCGGG	1560		
QY	601	GCAATGAGGCTCGGGTGGGAGGATGGCTCTGTACCCCGGCGTCCCGGCTTAGTT	660	Db	1501	TCACCGGGGGAAGTGGCGGCCGCCACCGCTGGGCTTGTGTGCTCTTACACGAGCG	1560		
Db	601	GCAATGAGGCTCGGGTGGGCGGATGGCTCTGTCTCCCGGTGGCTCTCGGCGTAGCT	660	QY	1561	CGTCTCAGAAATCCAGCTTGTGAATACCAACGCGAGCTGGCATCAACAGGACTTGCCC	1620		
QY	661	GGGGCCCCACGACCCCGGCTAGGTCGCTTAAGGTAAAGTCAATCATGATACCTTTA	720	Db	1561	CCAAGCAGAAATCCAACTGATCAACACCAACGCGAGTGGCATCAATAGCAGCGCT	1620		
Db	661	GGGGCCCCACGACCCCGGCTAGGTCGCTTAAGGTAAAGTCAATCATGATACCTTTA	720	QY	1621	TAAATTCGAATGACTCCCTCCAACTGGGTCTTTGCGCGGCTGTTTACGCAACAAGT	1680		
QY	721	CATGGGCTTCGCGATCTCATGGGTATCATTCGCTCGTGGCGCCCGCTTAGGGGCG	780	Db	1621	TGAATCTCAATGAAGCCTTAAACCGGCTGGTGTAGCAGGGCTCTTCTATCAGCACAAAT	1680		
Db	721	CATGGGCTTCGCGATCTCATGGGTATCATTCGCTCGTGGCGCCCGCTTAGGGGCG	780	QY	1681	TCAACTCGTCCGGGTGCCGAGCGCATGGCCAGCTGCCGCCCAATGACTGGTTCGCCC	1740		
QY	781	CTGCGAGGCTTGGCACACGCTTCCGGGTTCTGGAGGAGCGGCTGAATATGCAACAG	840	Db	1681	TCAACTCTTTCAGGCTGCTCTGAGAGGTGGCCAGCTGCCAGCGCTTACCAGTATTTGCC	1740		
Db	781	CTGCGAGGCTTGGCGCATGGGTCGCGGTTCTGGAAGACGGCTGAATATGCAACAG	840	QY	1741	AGGGTGGGGCCCCATACCTATATCTAAGCCTAATCAGCTCGGATCAGAGGCTTATTTGCT	1800		
QY	841	GGAACTTGGCCGGTGTCTTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900	Db	1741	AGGGCTGGGGTCTTATCAGTTATGCCAAACGGAAGCGGCTCGACGAAACGCGCTTACTGCT	1800		
Db	841	GGAACTTCTCTGTTGCTCTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900	QY	1801	GGCATTAGCGGCTTCGACCGGTGTGTGCTGTACCCGCGTCCGAGGTGTGCTCCAGTGT	1860		
QY	901	TCCAGCTTCGCTTATGAAGTGGCAAGGTGTCCGGGATATACCATGTACGAACGACT	960	Db	1801	GGCATTAGCGGCTTCGACCGGTGTGTGCTGTACCCGCGTCCGAGGTGTGCTCCAGTGT	1860		
Db	901	TCCAGCTTCGCTTATGAAGTGGCAAGGTGTCCGGGATATACCATGTACGAACGACT	960	QY	1861	ATTCGTTTACCCCAAGCCCTGTTGTTGGTGGGACCAACCGATGTTTCCGGTGTCCCTACGT	1920		
				Db	1861	ATTCGTTTACCCCAAGCCCTGTTGTTGGTGGGACCAACCGATGTTTCCGGTGTCCCTACGT	1920		
				QY	1921	ATAGCTGGGGGAGAAATGAGACAGAGTGTCTCTCAACAACAACGCGCTCCGCCACAAG	1980		
				Db	1921	ACAGCTGGGGTGCATAATGATAGGATGCTTCTGCTCTTAAACAACACGAGCCACCGCTGG	1980		
				QY	1981	GCAACTGGTTCGGCTGTATCGATGAATAGTACTGGGTTCACTAAGACGTGCGGAGGTC	2040		
				Db	1981	GCAATTTGGTTCGGTGTACTTGGATGAATCACTGGATTCACCAAAAGTGTGCGGAGCGC	2040		

4201	Db	CTGGCAGCCCACTCAGCTACTCCACCTACGGCAAGTTCTTGGCCAGCGCGGTGCTCAG	4260
4261	Qy	GGGGCGCCTATGACATCAATAATGTGATGTCACACTCAACTGACTCGACTACCATCT	4320
4261	Db	GAGTGCTTATGACATAAATTTGTGACAGGTGCCACTCCACGGATGCCATCATCT	4320
4321	Qy	TGGGCATCGGCACAGTCTCTGGACCAACGCGGAGCGGCTGGAGCGCGGCTCGTGTGCTCG	4380
4321	Db	TGGGCATCGGCACGTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTGTGCTCG	4380
4381	Qy	CCACCGCTACACTCTCGGGATCGGTTACGTTGCGCACACCCCAATATTCGAGGAATAGGCC	4440
4381	Db	CCACTGCTACCCCTCGGGCTCGCTCACTGTGTGCCATCTTACATTCGAGGAGGTGCTC	4440
4441	Qy	TGTCCAACAAATGAGAGATCCCTTCTATGGCAAAAGCCATCCCAATGTGAGGCCATCAAGG	4500
4441	Db	TGTCACACCACGGAGAGATCCCTTTTACGGCAAGGTATCCCCCTCGAGGTGATCAAGG	4500
4501	Qy	GGGGGAGGCATCTCATTTTCTGCGCATTCCAAGAAAGAAATGTGACGAGCTCGCCGCAAGC	4560
4501	Db	GGGGAGACATCTCATCTTCTGCCACTCAAAGAAAGAAAGTGCAGCAGAGCTCGCCGCAAGC	4560
4561	Qy	TGACAGCCTCGGACTGAACGCTGTGACATATTACCGGGCCCTTGATGTGTCGTCAATAC	4620
4561	Db	TGTCGCGAATTGGGCATCAATTCGCGTGGCCCTACTACCGCGGTCTTGACGTGTCTGTCATCC	4620
4621	Qy	CGCCTATCGGAGACGTGTGTGTCGCAACAGACGCTCTAATGACGGGTTTCACCGCGG	4680
4621	Db	CGACCAAGCGCGATGTGTGCTGTGTCGACCGATGCTCTCATGCTGGCTTTACCGGCG	4680
4681	Qy	ATTTGACTCAGTGTATCGACTGCAATACATGTGTCAACCACAGACGTGCACTTCAGCTTGG	4740
4681	Db	ACTTGGACTCTGTGATAGACTGCAACACGCTGTGCTCACTCAGACAGTCGATTCAGGCTTG	4740
4741	Qy	ATCCACCTTCAACATTTGAGACGACGACCGTGCGCCCAAGACGGGTGTGCGCTCGCAAC	4800
4741	Db	ACCTACCTTTACCAATTTAGACAACACACGCTCCGCCAGGATGCTGTCTCCAGGACTCAAC	4800
4801	Qy	GGCGAGTGAACCTGGCAGGGGTAGGAGTGGCATCTACAGTTTGTGACTCCAGGAGAAC	4860
4801	Db	GCCGGGCGAGACTTGGCAGGGGGAAGCAGGCATCTACAGATTTGTGGCACCGGGGAGC	4860
4861	Qy	GGCCCTCGGCATGTTCCGATCTTCGGTCTCTGTGAGTGCATAGACGGGGCTGTGCTT	4920
4861	Db	GCCCTCGGCATGTTGCACTCGTCCGCTCTCTGTGAGTGTATGACGCGGGCTGTGCTT	4920
4921	Qy	GGTATGAGTCAACGCCCGCTGAGACCTCGGTTAGTTTGGGGCTTACTAAATACACCAG	4980
4921	Db	GGTATGAGCTCACGCCCGCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGG	4980
4981	Qy	GGTTGCCGCTGCGCAGGACCATCTGAGTTCCTGGGAGAGCGTCTTCACAGGGCTCAACC	5040
4981	Db	GGCTTCGCGTGTGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGGCTCACTC	5040
5041	Qy	ACATAGATGCCACTTCTCTGTCCAGACTAAACAGGCAGGAGACAACTTTCTTCACTCGG	5100
5041	Db	ATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACTTTCTTACCTCGG	5100
5101	Qy	TGGCATATCAAGCTACAGTGTGCCCGCAGGGCTCAAGTTCACCTCCATGTGTGGGACCAA	5160
5101	Db	TAGCGTACCAAGCCACCGTGTGCGCTAGGGCTCAAAGCCCTCCGCCCATCGTGGGACCA	5160
5161	Qy	TGTGGAAGTGTCTATACGGCTGAACCTACACTGTGACGGGCCAACCCCTGCTGTATATA	5220
5161	Db	TGTGGAAGTGTGTGATCGCCCTTAAACCCACCCCTCCATTGGGCCAACCCCTGCTATACA	5220
5221	Qy	GGCTAGGAGCGGTGCAAAATCAGGTCTATCTTCACACACCCCATAACTAAATACATATGG	5280
5221	Db	GACTGGGGCGCTGTTTCAGNATGAGTTCACCTTGACGCAACCCCATCACCAATACATCATGA	5280
5281	Qy	CATGCATGTCGGCTGACCTGGAGGTGCTGCTACTAGCACCTGGGTGCTGTGTAGGCGGAGTCC	5340

Db	5281	CATGCATGTGCGGCCGACACTGGAGAGTTCGTACAGACACCTGGGTGTCGTGTGGCGGGCTCC	5344
Qy	5341	TTCCAGCTTTGGCCGCGATACTGCCTCAGCACAGGCAGTGTGGTCAATTGTGGCAGAGATCA	5400
Db	5341	TGCGTCTCTGGCCGCGTATTGCTGTCAACAGGCTGCGTGTGTATAGTGTGGCAGGATTG	5400
Qy	5401	TCTTGTCCGGGAAGCAGCAGTGTCTGTTCCGACAGGGAAGTCTCTTACCAGGAGTTCGATG	5460
Db	5401	TCTTGTCCGGGAAGCCGCAATTATACCTGTACAGGAGGTTCTCTACCAGGAGTTCGATG	5460
Qy	5461	AGATGAAGAGTGTGCTCTACAACCTTCTTTACATCAGCAGAGGAATTCAGCTCGCCGAGC	5520
Db	5461	AGATGGAGAGTGTCTCTCAGCACTTACCGTACATCCAGCAAGGATGATGCTCGCTGAGC	5520
Qy	5521	AATTCAAGCAAAAGGCGCTCGGTTGTTCGCAACGGCCCAACGAAGCAGCGAGGCTGTCTG	5580
Db	5521	AGTTCAAGCAGAAGGCCCTCGGCTCTCTGAGACCCGCTCCCGCCCAAGCAGAGGTTATCA	5580
Qy	5581	CTCCGTGTGTGGAGTCCAGTGGCGAGCCTTTCAGACCTTCTGGGCGAAGACATCTGGA	5640
Db	5581	CCCTGTGTCTCCAGACCACTGGCAGCAAACTCGAGGTCTTCTGGGCGAAGACATCTGGA	5640
Qy	5641	ATTTCATCAGCGAAATACGTACTATAGCAGCGTTATCCACTCTGCTCTGGAAAAACCCGCGA	5700
Db	5641	ATTTTCATCAGTGGATACATATCTTGGCGGCTGTCTCAGCTGCTGTGTAAACCCGCGCA	5700
Qy	5701	TAGCATCATTGATGGCATTTTACAGCTTCTATCTATAGCCGCTCACACCCAAAAACACC	5760
Db	5701	TTGCTTCATTGATGGCTTTTACAGCTGCCGTCCACAGCCCACTTAACCACTTGGCGCAACCC	5760
Qy	5761	TCTGTTTTACATCTTGGGGGGATGGTGGCTGCCCACTTCGCTCTCCACAGCGCTGCGT	5820
Db	5761	TCCTCTTCAACATATTGGGGGGGTGGTGTGCTGCCAGCTTCGCGCCCCCGGTGCGGCTA	5820
Qy	5821	CAGCTTTCTGGGCGCGGCATCCGCGAGCGGTGTGTGGCAGCATAGGCGCTTTGGGAAGG	5880
Db	5821	CCGCTTTTGTGGCGCTGGCTTAGCTGGCGCGGCATCCGACGGTGTGGACTGGGGAAGG	5880
Qy	5881	TGCTCGTGCACATCTTGGCGGGCTATGGGCGAGGGTGTAGCCGCGCATCTCGTGGCCTTTA	5940
Db	5881	TCCTCGTGGACATTTCTTGACAGGTATTGGCGGGCGTGGCGGAGCTCTTGTAGCCTTCA	5940
Qy	5941	AGTTCATAGCGGGGAGGTGCCCTCCACCGAGACCTGGTCAACTTACTCCCTGCCATCC	6000
Db	5941	AGATCATGAGCGGTGAGGTCCCTCCACGAGGACCTGGTCAATCTGCTGCCCGCCATCC	6000
Qy	6001	TCTCTCTGTGGCTCGTCTGTGGGTGTGTGTGCGCAGCAATACTTCGCTCGGCACGTGG	6060
Db	6001	TCTCGCTGTGAGCCCTTTGTAGTCCGTGTGTCTGCGCAGCAATACTTCGCGCGGCACGTTG	6060
Qy	6061	GCCCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGTAGCGTTCGCTTCGCGGGGTA	6120
Db	6061	GCCCGGCGAGGGGCGCATGCAATGGATGAACCGGCTTAATAGCTTCGCTCCCGGGGGA	6120
Qy	6121	ACCAAGTCTCCCTTAGCACAATGTGCTGTAGAGCGACGTGTGACAGCACTGTCACTCAGA	6180
Db	6121	ACCATGTTTCCCCACGCACTACGTGCGGAGAGCGATGACAGCCGCGCGGTCACTGCCA	6180
Qy	6181	TCCTCTCTAGCCTTACCATCACTCAACTGCTGAAGCGGCTCCACCACTGGATTAATAGG	6240
Db	6181	TACTCAGCAGCCTCACTGTAAACCAAGCTTCTGAGCGCACTGCATCAGTGGATAGCTCGG	6240
Qy	6241	ACTGCTCTACGCCATGTCTCGGCTCGTGGCTTAAGGATGTTTGGATTTGGATATGCAAGG	6300
Db	6241	AGTGTAACCACTCCATGCTCCGGTTCCTGGCTTAAGGGACATCTGGGCACTGGATATGCGAGG	6300
Qy	6301	TGTTGACTGACTTCAAGACCTTGGCTCCAGTCCAACTCTGCGCGGGTTACCGGGAGTCC	6360
Db	6301	TGCTGAGCGACTTTTAAGACCTTGGCTGTAAAGCACAAGCTCATGTCCAACAATGCCTCGGATTC	6360
Qy	6361	CTTTCCTGTCAATGCCACGCGGTTACAGGGAGTCTGCGGGGGGACGGCATCATGCAAA	6420
Db	6361	CCTTTGTGTCTGCGCAGCGCGGTTATAGGGGGGTTCTGGCGAGGAGACGGCATTTATGCACA	6420

Db	8581	CGGGGTTCCAGGAGGACGGCGGAGCGTCTGAGAGCCTTTCAGGAGGCTATGACAGGTA	CT	CGGGGTTCCAGGAGGACGGCGGAGCGTCTGAGAGCCTTTCAGGAGGCTATGACAGGTA	8640
Qy	8638	CCGCCCCCGGGGATCCGCCCAACCAAGATACGACCTGGAGCTGATAAATCATCATGTT	CT	CCGCCCCCGGGGATCCGCCCAACCAAGATACGACCTGGAGCTGATAAATCATCATGTT	8697
Db	8641	CCGCCCCCGGGGACCCCCCAACCAAGATACGACCTGGAGCTGATAAATCATCATGCT	CT	CCGCCCCCGGGGACCCCCCAACCAAGATACGACCTGGAGCTGATAAATCATCATGCT	8700
Qy	8698	CCTCCAATGTGTGAGTGGCGCACGATCTCGCAAAAGGTTATTA	CT	CCTCCAATGTGTGAGTGGCGCACGATCTCGCAAAAGGTTATTA	8757
Db	8701	CCTCCAAGTGTGAGTGGCGCACGAGCGGCTGGAAAGAGGGTCTACTACCTTTACCCGCTG	CT	CCTCCAAGTGTGAGTGGCGCACGAGCGGCTGGAAAGAGGGTCTACTACCTTTACCCGCTG	8760
Qy	8758	ACCCACACCCCTTGACAGGGCTGGTGGGAGACGCTAGACACATCCCAATCAACT	CT	ACCCACACCCCTTGACAGGGCTGGTGGGAGACGCTAGACACATCCCAATCAACT	8817
Db	8761	ACCCACACCCCTTGACAGGGCTGGTGGGAGACGAGCAACACATCTCCAGTCAAT	CT	ACCCACACCCCTTGACAGGGCTGGTGGGAGACGAGCAACACATCTCCAGTCAAT	8820
Qy	8818	CTTGGCTAGGCAATATCATCATGTATGCGGCCCACTATGGGCAAGGATGATTCTGATGA	CT	CTTGGCTAGGCAATATCATCATGTATGCGGCCCACTATGGGCAAGGATGATTCTGATGA	8877
Db	8821	CTTGGCTAGGCAATATCATCATGTGTCCTCCACCTGTGGCGAGGATGATCTGATGA	CT	CTTGGCTAGGCAATATCATCATGTGTCCTCCACCTGTGGCGAGGATGATCTGATGA	8880
Qy	8878	CTCACTTTTCTCCATCTTCTAGCTCAAGAGCACTTGAAAAAGCCCTGGATTGTGAGA	CT	CTCACTTTTCTCCATCTTCTAGCTCAAGAGCACTTGAAAAAGCCCTGGATTGTGAGA	8937
Db	8881	CCCAATTTCTTAGCGTCTCATAGCCAGGGATCAGCTTGAACTGAGGCTCTTAACTGTGAGA	CT	CCCAATTTCTTAGCGTCTCATAGCCAGGGATCAGCTTGAACTGAGGCTCTTAACTGTGAGA	8940
Qy	8938	TCTACGGGCTTGCTACTCCATTGAGGCACTTGACCTACCTCAGATCAATTTGAAGACTCC	CT	TCTACGGGCTTGCTACTCCATTGAGGCACTTGACCTACCTCAGATCAATTTGAAGACTCC	8997
Db	8941	TCTACGGAGGCTGCTACTCCATAGAACCACTGGATCTACTCCAAATCAITTCAAAGACTCC	CT	TCTACGGAGGCTGCTACTCCATAGAACCACTGGATCTACTCCAAATCAITTCAAAGACTCC	9000
Qy	8998	ATGGTCTTAGCGCAATTTACACTCCACAGTTTACTCTCCAGGTGAGATCAATAGGGTGGCTT	CT	ATGGTCTTAGCGCAATTTACACTCCACAGTTTACTCTCCAGGTGAGATCAATAGGGTGGCTT	9057
Db	9001	ATGGCTCAGCGCATTTTCACTCCACAGTTTACTCTCCAGGTGAAATCAATAGGGTGGCG	CT	ATGGCTCAGCGCATTTTCACTCCACAGTTTACTCTCCAGGTGAAATCAATAGGGTGGCG	9060
Qy	9058	CATGCCTCAGGAAACTTTGGGGTACCACTCTGGGAACTCGGAGACATCGGGCCAGAGTG	CT	CATGCCTCAGGAAACTTTGGGGTACCACTCTGGGAACTCGGAGACATCGGGCCAGAGTG	9117
Db	9061	CATGCCTCAGAAAACTTTGGGGTCCGGCTTCGCGAGCTTGCGAGACACCGGGCCGGAGCG	CT	CATGCCTCAGAAAACTTTGGGGTCCGGCTTCGCGAGCTTGCGAGACACCGGGCCGGAGCG	9120
Qy	9118	TCCGGCTTAGCTACTCTCCAGGGGGGAGGGCCGCACTTGTTGGCAGATACCTCTTTA	CT	TCCGGCTTAGCTACTCTCCAGGGGGGAGGGCCGCACTTGTTGGCAGATACCTCTTTA	9177
Db	9121	TCCGGCTTAGGCTTCTGTTCAGAGGAGCGAGGCTGCCATATGTGGCAAGTACTCTTCA	CT	TCCGGCTTAGGCTTCTGTTCAGAGGAGCGAGGCTGCCATATGTGGCAAGTACTCTTCA	9180
Qy	9178	ACTGGGAGTAGAAGCAACAGCTTAAACTCACTCCAAATCCGGCCCGCTCCAGCTGGACT	CT	ACTGGGAGTAGAAGCAACAGCTTAAACTCACTCCAAATCCGGCCCGCTCCAGCTGGACT	9237
Db	9181	ACTGGGAGTAGAAGCAACAGCTCAACTCACTCCAAATAGCGCCGCTGGCGCTGACT	CT	ACTGGGAGTAGAAGCAACAGCTCAACTCACTCCAAATAGCGCCGCTGGCGCTGACT	9240
Qy	9238	TGCTGGCTGGTTCTGCTCGCTGGTTACAGCGGGGGAGACATATATCAGAGCTGTCTCGT	CT	TGCTGGCTGGTTCTGCTCGCTGGTTACAGCGGGGGAGACATATATCAGAGCTGTCTCGT	9297
Db	9241	TGTCGGTGTGTTACGCGCTGGCTACAGCGGGGGAGACATTTATCAGACGGTGTCTCATG	CT	TGTCGGTGTGTTACGCGCTGGCTACAGCGGGGGAGACATTTATCAGACGGTGTCTCATG	9300
Qy	9298	CCGACCCCGCTGTTCCGTTGTGCGCTACTCTCACTCTCTGTAGGGGTAGGCAATTTACC	CT	CCGACCCCGCTGTTCCGTTGTGCGCTACTCTCACTCTCTGTAGGGGTAGGCAATTTACC	9357
Db	9301	CCGCGCCCGCTGTTCTGGTTTGGCTACTCTGCTCGCTGCGAGGGGTAGGCAATTTACC	CT	CCGCGCCCGCTGTTCTGGTTTGGCTACTCTGCTCGCTGCGAGGGGTAGGCAATTTACC	9360
Qy	9358	TGCTCCCCAAACCGATGAACGGGAGCTAAACCACTCCAGGCCTTTAAGCCATTTCCCTGTT	CT	TGCTCCCCAAACCGATGAACGGGAGCTAAACCACTCCAGGCCTTTAAGCCATTTCCCTGTT	9416
Db	9361	TCCCTCCCAACCGATGAAGTGTGGGTAAACACTCCGGCTCTTTAGGCCATTTCTCTGTT	CT	TCCCTCCCAACCGATGAAGTGTGGGTAAACACTCCGGCTCTTTAGGCCATTTCTCTGTT	9420
Qy	9417	TT	CT	TT	9476
Db	9421	TT	CT	TT	9480
Qy	9477	TT	CT	TT	9487
Db	9481	TT	CT	TT	9540
Qy	9488	CTTCTTTAATGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAAAGGTCGGTGAG	CT	CTTCTTTAATGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAAAGGTCGGTGAG	9547
Db	9541	CTTCTTTAATGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAAAGGTCGGTGAG	CT	CTTCTTTAATGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAAAGGTCGGTGAG	9600
Qy	9548	CCGCAATGACTGCAGAGATGTGTATCTGTGGGCTCTCTGCAGATCATGT	CT	CCGCAATGACTGCAGAGATGTGTATCTGTGGGCTCTCTGCAGATCATGT	9595
Db	9601	CCGCAATGACTGCAGAGATGTGTATCTGTGGGCTCTCTGCAGATCATGT	CT	CCGCAATGACTGCAGAGATGTGTATCTGTGGGCTCTCTGCAGATCATGT	9648

RESULT 11

```

US-09-917-563-5
; Sequence 5, Application US/09917563
; Publication No. US20030073080A1
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; TYPE OF INVENTION: VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/917,563
; FILING DATE: 27-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/238,076
; FILING DATE: 26-JAN-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-917-563-5

Query Match      65.3%; Score 6264; DB 3; Length 12980
Best Local Similarity 78.4%; Pred. No. 0;
Matches 7565; Conservative 0; Mismatches 2030; Indels 5

QY      1      GCCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCCCTGTGAGG
DB      1      GCCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCCCTGTGAG
QY      61     TC TTCACGACGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGCGTGCGAGGAGG
DB      61     TC TTCACGACGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGCGTGCGAGGAGG
QY      121    CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGG
DB      121    CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGG
QY      181    GACGACCGGGTCCCTTCTTTGGATCAACCCCGTCAATGCTCGAGATTGGGTCGAG
DB      181    GACGACCGGGTCCCTTCTTTGGATCAACCCCGTCAATGCTCGAGATTGGGTCGAG
QY      241    GCGAGACTGCTAGCCGAGTAGTGTGGGTGCGGAAGGCCCTTGTGGTACTCT
DB      241    GCAAGACTGCTAGCCGAGTAGTGTGGGTGCGGAAGGCCCTTGTGGTACTCT
QY      301    GTGCTTGCGAGTGCCCCCGGAGGTCTCTAGACCGTGACCATGACGACG
DB      301    GTGCTTGCGAGTGCCCCCGGAGGTCTCTAGACCGTGACCATGACGACG

```

QY 361 CTCAAAGAAAAACCAACGTAACACAAACCGCGGCCCAACAGACGTCAGATTCCCGGCG 420
DB |||||
QY 361 CTCAAAGAAAAACCAACGTAACACAAACCGTCGCCCAACAGACGTCAGATTCCCGGCG 420
DB |||||
QY 421 GTGGTCAGATCGTTGGTGGAGTTTACCTGTGTGCGCGCAGGGGCCCAAGTTGGGTGTC 480
DB |||||
QY 421 GCGGTTCAGATCGTTGGTGGAGTTTACTTGTGTGCGCGCAGGGGCCCTAGATTGGGTGTC 480
DB |||||
QY 481 GGGGACTAGGAGGCTTCCGAGCGGTCCGAACCTTCGTGGAGGCGCAACCTATCCCAA 540
DB |||||
QY 481 GCGGACGAGGAAGACTTCCGAGCGGTCCGAACCTTCGAGGTTAGAGCTCAGCTATCCCA 540
DB |||||
QY 541 AGGCTCGCGCACCCGAGGCGAGGGCTGGGGTTCAGCCCGGTTACCTTTGGCCCTCTATG 600
DB |||||
QY 541 AGGACAGTTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCTTTGGCCCTCTATG 600
DB |||||
QY 601 GCAATGAGGGCTGGGGTGGGAGGATGGCTCTGTCAACCGCGGCTCCGGGCTAGTT 660
DB |||||
QY 661 GGGGCCCCACGACCCCGGCTAGTTCGGTAACTTTGGTTAAGTTCATCGATACCCCTTA 720
DB |||||
QY 661 GGGGCCCCACGACCCCGGCGTAGGTTCGGCAATTTGGGTAAGTTCATCGATACCCCTTA 720
DB |||||
QY 721 CATGGGCTTCGCGCATCTCATGGGTTACATTCGCTCGTCGGCGCCCGCTAGGGGCG 780
DB |||||
QY 721 CGTGGGCTTCGCCGACCTCATGGGTTACATACCGCTCGTCGGCGCCCGCTCTTGAGGCG 780
DB |||||
QY 781 CTGCCAGGCGCTTGGCACACGGTGTCCGGGTTCTGGAGACGGCTGAATATGCAACAG 840
DB |||||
QY 781 CTGCCAGGCGCTTGGCGCATGCGGTTCCGGTTCTGGAGAGCGGCTGAATATGCAACAG 840
DB |||||
QY 841 GGAACTTGCCCGGTTGCTCTTCTCTATCTCTCTCTGCTGCTGCTGCTGCTGCTGTTGACCA 900
DB |||||
QY 841 GGAACTTCTCTGTTGCTCTTCTCTATCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 900
DB |||||
QY 901 TCCGAGCTTCGCTTATGAAGTGGCAACGCTGTCGGGATATACCATGTTCAGCAACGACT 960
DB |||||
QY 901 TCCGCGCTTCAGCTACCAAGTGGCAATTCCTCGGGCTTTTACCATGTTCACCAATGAT 960
DB |||||
QY 961 GCTCAACTCAAGCATTTGTATGAGGACGAGCGATGATCATATGATATCTCCGGGTGCG 1020
DB |||||
QY 961 GCGCTAACTCGAGTATTGTACGAGGCGCGGATGCCATCTCGCACCTCCGGGGTGTG 1020
DB |||||
QY 1021 TCGCCTGTGTTCAGAGGGTAACAGCTCCGTTGCTGGGTAGCGCTCACTCCACGCTCG 1080
DB |||||
QY 1021 TCCCTTGGCTTCGCGAGGGTAACGCTCGAGGTGTGGGTGGCGGTGACCCCGACGCTGG 1080
DB |||||
QY 1081 CGGCGAGGAATGCCAGCGTCCCACCTACGACAATACGAGCGCACGTCGACTTGTCTGTTG 1140
DB |||||
QY 1081 CCACAGGAGCGGCAAACTCCCAACACGAGCTTCGAGCTCATATCGATCTGCTTGTGCG 1140
DB |||||
QY 1141 GGACGCTCTTCTGCTCCGCTATGTAGTGGGGATCTCTCGGATCTATTTTCTCG 1200
DB |||||
QY 1141 GGAGCGCCACCTCTGCTCGGCCCTCTAGTGGGGACCTGTGCGGGTCTGCTTCTTG 1200
DB |||||
QY 1201 TCTCCAGCTGTTACCTTCTCGCTTCGCGGATGAGACAGTGCAGGACTCGAACTGCT 1260
DB |||||
QY 1201 TTGGTCAACTGTTTACCTTCTCTCCAGGCGGCACTGGACGAGCAAGACTGCAATTTGT 1260
DB |||||
QY 1261 CAATCTATCCCGGCATGATCAGGTACCGGATGCTTGGGATGATGATGAACTGCT 1320
DB |||||
QY 1261 CTATCTATCCCGGCATATAACGGGTATCGATGGCATGGGATGATGATGAACTGCT 1320
DB |||||
QY 1321 CACCTACAAACAGCCCTAGTGGTGTGCGAGTTGTCTCGGATCCCAACAGCTGCTGGGACA 1380
DB |||||
QY 1321 CCCCTACGCGAGGTTGGTGGTAGCTCAGCTGCTCGGATCCCAACAGCCATCATGGACA 1380
DB |||||
QY 1381 TGGTGGCGGGGCCCACTGGGAGTCTCGGCGGCTTCGCTACTATTCATGTTGAGGA 1440
DB |||||
QY 1381 TGATCGCTGTGTCTAGTGGGAGTCTCGCGGCAATAGCTGTGATGATGATGATGATGATGATG 1440

QY 1441 ACTGGCTAAGGTTCTGATTGTGGCGTACTCTTTTGGCGGTTGACGGGGAGACCCACA 1500
DB |||||
QY 1441 ACTGGCGAAGGTCCTGTGTGCTGCTGCTATTTTGGCGGTCGACGCGGAAACCCACG 1500
DB |||||
QY 1501 CGACGGGAGGTTGGCGGCCACACCACTCCGGGTTACGTCCCTTTTCTCATCTGGGG 1560
DB |||||
QY 1501 TCACCGGGGAAGTGGCGGCCGACCAACGCTGGGCTTGTGGTCTCTCTACACAGGCG 1560
DB |||||
QY 1561 CGTCTCAGNAATCCAGCTTGTGAATACCAACGCGAGCTGGCACATCAACAGGAGTGGCC 1620
DB |||||
QY 1561 CCAAGCAGAACATCAACTGATCAACCAACGCGAGTTGSCACATCAATAGCAGGCT 1620
DB |||||
QY 1621 TAAATTGCAATGACTCCCTCCAACTGGGTTCTTTGGCGGCTGTFTTTCGACACAAAGT 1680
DB |||||
QY 1621 TGAATGCAATGAAAGCTTAAACACCGCTGTTAGCAGGCTCTTCTTATCAGCACAAAT 1680
DB |||||
QY 1681 TCAACTCGTCCGGTGGCGGAGCGCATGGCAGCTGGCGGCCCACTTGACTTGTTGGGCC 1740
DB |||||
QY 1681 TCAACTCTTCAGGCTGTCTGAGAGGTTGGCAGCTGCCGAGCGCTTACCAGATTGCGCC 1740
DB |||||
QY 1741 AGGGTGGGGCCCATCACCTATATAAGCTTAACAGCTCCGATCAGAGGCTTATTGCT 1800
DB |||||
QY 1741 AGGCTGGGGTCTTATCAGTTATGCCAACGAGGGCTCGACGAACGCCCTACTGCT 1800
DB |||||
QY 1801 GGCATTAGCGCCCTCGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
DB |||||
QY 1801 GGCATTACCTCCAAAGACCTTGTGGCATTTGTCCCGCAAGAGCGTGTGTGTGTGTGTGT 1860
DB |||||
QY 1861 ATTGTTTACCCCAAGCCCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
DB |||||
QY 1861 ATTGTTTCACTCCAGCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
DB |||||
QY 1921 ATAGTGGGGGAGATGAGACAGCGTGTGTCTTCAACAAACACGCGTCCGCCACAAG 1980
DB |||||
QY 1921 ACAGTGGGGTGCAATGATACGGATGTCTTGTCTTAAACAAACAGGCGCACCGCTGG 1980
DB |||||
QY 1981 GCAACTGCTTCGGCTGTACATGGATGAATAGTACTGCGGTTCACATAAGACGTGCGAGGTC 2040
DB |||||
QY 1981 GCAATTGCTTCGGTTGTACCTGGATGAATCAACTGGATTCAACAAAGTGTGCGGAGCGC 2040
DB |||||
QY 2041 CCCGTTGTAAATCGGGGGGTCCGTAAACCGACCTTGATCTGCCCCACGGACTGCTTCC 2100
DB |||||
QY 2041 CCCCTTGTGTATCGGAGGGTGGGCAACACACCTTGTCTCTGCCCCCATGATTTTTC 2100
DB |||||
QY 2101 GGAGCACCCGAGGCTTACTACAAATGTGGCTCGGGCCCTGTTGTGACCTAGT 2160
DB |||||
QY 2101 GCAAGCATCCGAGACACACTCTCGGTGGGCTCCGGTCCCGTCCCGTCCCGTCCCGTCC 2160
DB |||||
QY 2161 GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCT 2220
DB |||||
QY 2161 GCATGGTCACTACCCGTATAGGCTTTGGCACTATCTTTGTACCATCAATTAACCATAT 2220
DB |||||
QY 2221 TTAAGTTAGGATGATGTGGGGGCGTGGAGCACAGGCTCAATGCGCATGCAATTTGA 2280
DB |||||
QY 2221 TCAAAAGTCAGGATGTACGTGGGAGGGTTCGAGCACAGGCTGGAAAGCGGCTTCAACTGA 2280
DB |||||
QY 2281 CTCGAGGAGAGCGGTGTAATTTGGAGACAGGATAGGTTCAGAACTCAGCCCGCTGCTGC 2340
DB |||||
QY 2281 CGCGGGCGAAGCGCTGTGATCTGGAAGACAGGACAGGCTCGAGCTCAGCCCTTGTCTGC 2340
DB |||||
QY 2341 TGTCTAACACAGAGTGGCAGATACTGCGCTGTGCTTTTCCACCCCTACCGGCTTTATCCA 2400
DB |||||
QY 2341 TGTCCACACACAGTGGGAGGCTCTTCCGTGTTCTTTCAAGACCTTCCAGGCTTTGTCCA 2400
DB |||||
QY 2401 CTGGTTTGTATCTCATCTCCATCAGAAACATCTGGGACGTGCAATACCTGTACGGTGTAGGT 2460
DB |||||
QY 2401 CGGGCTCATCCACCTCCACAGAACATTTGTGGACGTGCAGTACTTGTACGGGGTAGGGT 2460
DB |||||
QY 2461 CAGCGTTGTCTCTTTTGCATCAATCGGAGTACATCTGTGTGTGTGTGTGTGTGTGTGTGT 2520
DB |||||
QY 2461 CAAGCATCGGCTCTGGGCCATTAAGTGGAGTACGTGTTCTCTCTGTGTGTGTGTGTGTGT 2520
DB |||||
QY 2521 CAGACGCGCGGTGTGTGCTGCTTGTGTGATGATGCTGTGTATGATGATGATGATGATGATG 2580

[illegible]

Db	3601	CGAGGACCATCGCATCACCCAAAGGTCCTGTCTCATCCAGATGTATTACCAATGTGGACCAAG	3660
Qy	3661	ACCTCGTCGGCTGGCAGGGCCCCCGGGCGCGCTCCATGACACCATGACGCTGTGGCA	3720
Db	3661	ACCTTGTGGCTGGCCCGCTCCTCAAGGTTCCCGCTCATTTGACACCCCTGCACCTCGCGCT	3720
Qy	3721	GCTCGGACCTTTACTTTGGTTCAGGACATGCTGATGTGTCATTCGGTGGCCGCGGAGGGG	3780
Db	3721	CCTCGGACCTTTACTCGTGGTTCAGAGGACGCGGATGTCAATTCCTGCGCGCGGAGGGT	3780
Qy	3781	ACAGCAGGGGAAGTCTACTCTCCCGCCAGGCCCCGTCTCTACCTGAAAGGCTCTCCGGGTG	3840
Db	3781	ATAGCAGGGGTAGCTTGTCTTCCGCCCGGCCCATTTCTTACTTGAAGGCTCCTCGGGG	3840
Qy	3841	GTCCATTTGCTTTGCCCTTCGGGGCACGTCGTGGGCGTCTTCGGGGTGTGTGTGCACCC	3900
Db	3841	GTCCGCTGTGTGCCCGCGGGACACGCCGTGGGCCCTATTTCAGGGCGCGGCTGTGCACCC	3900
Qy	3901	GGGGGTGCGGAAGGGGTGGACTTCATACCCGTTGAGTCTATGGAAATACCAATCGGCT	3960
Db	3901	GTGGAGTGGCTAAGGGGTGGACTTTATCCCTGTGGAAACCTTAGAGACAACCATGAGAT	3960
Qy	3961	CTCCGCTCTTCAGACAACTCAACCCCCCGGCTGTACCGCAGACATTTCCAAGTGGCAC	4020
Db	3961	CCCCGTTTTCAGGACAACTCTCTTCACACGAGTGCCTCCAGAGCTTCCAGTGGCCC	4020
Qy	4021	ATCTGCACGCTCTACTGGCAGCGGCAAGACACCAAAGTGCCTGCGGTATGACAGCCC	4080
Db	4021	ACCTGCATGCTCCACCGGCAGCGTTAAGAGCACCAAGGTCCCGGCTGCCTACGCAAGCCC	4080
Qy	4081	AAGGTTACAAGGTGCTCGTCTGAAACCCGTCCTGTCGGCCACCTTAGGTTTGGGGCGT	4140
Db	4081	AGGGCTACAAGGTGTTGGTGCTCAACCCCTCTGTGTGTCAAACGCTGGGCTTTGGTGCTT	4140
Qy	4141	ATATGTCGAAGGCACACGGTATCGACCTTAACTCAGAACTGGGGTAAAGACCAATTACCA	4200
Db	4141	ACATGTCGAAGGCCATGGGTTGATCTTAATATCAGACCGGGTCAGAACAATTACCA	4200
Qy	4201	CGGGCGGCTCCATTCGACTCCACCTATGGCAAGTTCCTTGGCGAGTGGCTGTCTGTG	4260
Db	4201	CTGSCAGCCCCATCAGCTACTCCACCTACGCGAAGTTCCTTGGCAGCGGGTGCTCAG	4260
Qy	4261	GGGGCGGCTATGACATCATTAATATGTGATGAGTGCCACTCAACTGACTCGACTACCATCT	4320
Db	4261	GAGGTGCTTTATGACATAAATAATTTGTGACAGGTGCCACTCCAGGATGCCACATCCATCT	4320
Qy	4321	TGGGCATCGGCACAGTCTCGACCAAGCGAGACCGCTGGAGCGGCTCGTGTGTCTCG	4380
Db	4321	TGGGCATCGGCACATGTCCTTGACACAGACTCGGGGGCGAGACTGGTTGTGTCTCG	4380
Qy	4381	CCACCGGCTACACCTCCGGGATCGGTACCGTGGCCACACCCCAATATCGAGGAATAGGCC	4440
Db	4381	CCAATGCTACCCCTCCGGGCTCGTCTACCTGTGTCCATCTCAATCGAGGAGGTGCTC	4440
Qy	4441	TGTCACAATATGAGAGATCCCTTCTATGGCAAGCCATCCCCATTTGAGGCCATCAAGG	4500
Db	4441	TGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAGGTGATCAAGG	4500
Qy	4501	GGGGGAGCATCTCATTTTCTGCCATTCGAAGAAATGTGACGAGCTCGCCCAAGC	4560
Db	4501	GGGGAAGACATCTCATCTTTCGCCACTCAAGAAGAAGTGGACGAGCTGCCCGGAGC	4560
Qy	4561	TGACAGGCGCTCGGACTGAACGCTGTAGCATATTACCGGGGCTTTGATGTGTTCGTATAC	4620
Db	4561	TGGTCGCAATTGGGCATCAATGCGCGTGCCCTACTACCGGGTCTTGACGTGTCTGTCTATCC	4620
Qy	4621	CGCCTATCGGAGAGCTGTGTGTCGTGGCAACAGAGCTCTAATGACCGGTTTACCCGGC	4680
Db	4621	CGACCAGCGCGATGTTGTCTGTCGTGTCGACCGATGCTCTCATGACTGGGCTTTACCGGGC	4680
Qy	4681	ATTTTGACTCAGTGCAGCTGCAATACATGTGTCACCGACAGTGCAGCTTCAGCTTGG	4740
Db	4681	ACTTCGACTCTGTGATAGACTGCAACACGTTGTCTACTCAGACAGTGCATTTAGSCCTTG	4740

Qy	4741	ATCCACCTTACCAATTGAGACGACGACCGTGCCCAAGAAGCGGTGTGCGCTCGCAAC	4800
Db	4741		
Qy	4801	GGCGAGTAGAACTGGCAGGGTAGGATGGCATCTACAGTTTGTGACTCCAGAGAAC	4860
Db	4801		
Qy	4861	GGCCCTCGGGCATGTTCAATCTTCGGTCTGTGTGAGTGTATGACGGGGCTGTGCTT	4920
Db	4861		
Qy	4921	GATATGAGTCAAGCCGCTGAGACCTCGGTTAGGTTCGGGCTTACCTAAATACACAG	4980
Db	4921		
Qy	4981	GTTTCCCGTCTGCCAGGACCATCTGGAGTTCTGGGAGAGGCTTTTCAAGGCTCACCC	5040
Db	4981		
Qy	5041	ACATAGATGCCACCTTCTGTCCAGACTAAACAGGCGAGAGCAACTTTCTTACCTGG	5100
Db	5041		
Qy	5101	TGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCATGCTGGGACCAAA	5160
Db	5101		
Qy	5161	TGTGGAAGTGTCTATACGGCTGAAACCTTACACTGCAGGGCGCAACACCCCTGCTGTATA	5220
Db	5161		
Qy	5221	GGCTAGGACCGTCCAAATAGAGTCTCTCACAACACCCCATAACTAAATACATCATGG	5280
Db	5221		
Qy	5281	CATGATGTGCGCTGACCTGGAGTGTGCTACCTAGCACCTGGTGTGCTGTAGGCGAGTCC	5340
Db	5281		
Qy	5341	TTGCAAGTTTGGCCGCTACTCCCTGAGACAGGAGTGTGCTCATTTGTGGCAGGATCA	5400
Db	5341		
Qy	5401	TCTTGTCCGGGAAGCCAGCTGTGTTCCCGACAGGGAAGTCTCTTACAGGAGTTTGGATG	5460
Db	5401		
Qy	5461	AGATGGAAGTGTGCTTCAACACTTCCCTTACATCGACGGGAATGCACTCGCCGAGC	5520
Db	5461		
Qy	5521	AAATCAAGCAAAAGCGCTCGGCTTTGTCAAACCGCCACCAAGCAAGCGGAGGCTGCTG	5580
Db	5521		
Qy	5581	CTCCCGTGTGAGTTCAAAGTGGGAGCCCTTTGAGACCTTTGAGACCTTTGGGCGAAGCATGTGGA	5640
Db	5581		
Qy	5641	ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGGAACCCCGGA	5700
Db	5641		
Qy	5701	TAGCATCATTTGATGGCATTTACAGTCTTCTATCACTAGCCCGCTCACCAACCCCAAAACACC	5760
Db	5701		
Qy	5761	TCCTGTTTAAACATTTTGGGGGAGTGGGTGGCTGCCCACTTCCCTCTCCAGCGTGGT	5820
Db	5761		

Qy	5821	CAGCTTTCTGTCGGCGCGGCATCGCCGAGCGGCTGTGGCAGCATAGGCTTTGGGAAGG	5880
Db	5821		
Qy	5881	TGCTGTGGACATCTTGGCGGGCTATGGGCGAGGGTAGCGGCGCACTCGTGGCCCTTTA	5940
Db	5881		
Qy	5941	AGGTCATGAGCGGAGGTGCCCTCCACCGAGGACCTTGGTCAACTTACTCCCTCCCATCC	6000
Db	5941		
Qy	6001	TCTCTCTGTCGCTCGCTCGGGCTCGTGTGCGCAGCAATACTCGTCGTCGCAACGTGG	6060
Db	6001		
Qy	6061	GCCCGGAGAGGGGCTGTGCAAGTGAATGAACCGGCTTAATAGCCTTCGCCCTCCCGGGGA	6120
Db	6061		
Qy	6121	ACCAGCTCTCCCTACGCACTATGTGCTGAGAGCGAGCTGCGAGCAGCTGTCACTCAGA	6180
Db	6121		
Qy	6181	TCCTCTCTAGCCTTACCCTCACTCAACTGTCTGAAGCGGCTCCACAGTGAATTAATGAGG	6240
Db	6181		
Qy	6241	ACTGCTCTACGCCATGTCTCGGCTCGTGGTGAAGGATGTTTGGGATTTGGATATGCACGG	6300
Db	6241		
Qy	6301	TGTTGACTGACTTCAAGACCTGGCTCCAGTCCAAACTCTCTGCCGGGTTCACGGGAGTCC	6360
Db	6301		
Qy	6361	CTTTCTCTCATGCAACCGGGGTACAAGGAGTGTGCGGGGGGACGGCATCATGCAAA	6420
Db	6361		
Qy	6421	CCACTGCCATGCGGAGCAAGATCCCGGACATGTCAAAAACGGTTTCCATGAGGATCG	6480
Db	6421		
Qy	6481	TAGGCGCTAGAACCTGCAAGCAACAGCTGGGCGAAGCTTCCCATCAACGCATACACCA	6540
Db	6481		
Qy	6541	CGGACCTTGCACACCTCCCGCGCGCCAACTATTCAGGGCGCTATGGCGGGTGGCTG	6600
Db	6541		
Qy	6601	CTGAGGAGTACGTGGAGTTACGCTGTGGGGATTTCCACTACGTGACGGGCAATGACCA	6660
Db	6601		
Qy	6661	CTGACAAAGTAAAGTGCCTCATGCCAGGTTCCGGCCCCCGAATTTCTCACGGAGGTGATG	6720
Db	6661		
Qy	6721	GAGTGGGCTGCACAGGTACGCTCCGGGTCAAAACCTCTTTACGGGAGGACGTCACGT	6780
Db	6721		
Qy	6781	TCCAGGTGGGCTCAACCAATPACTTGGTGGGTGCGAGCTCCCATTCGAGGCCGACCGG	6840
Db	6781		
Qy	6841	ACGTAAAGTGTCTTACTTCCATGTCTACCGATCCCTCCCATTTACAGCAGAGCGGCTA	6900
Db	6841		
Qy	6901	AGCGTAGGCTGGCTAGAGGGTCTCCCGCTCTTTTAGCCAGCTCATCAGTACGCCAGTTGT	6960
Db	6901		

Db	6901	GGAGAGGTTGGCGAGAGGTCACCCCTTCTATGCGCAGCTCCTCGGCAGCAGCTGT	6960
Qy	6961	CTCGCTCTCTTTGAAGGCGACATGACATACCCACCATGACTCCCGGAGCGTGAACCTCA	7020
Db	6961	CGCTCCATCTCTCAAGGCAACTTGCACCGCAACATGACTCCCTCGAGCGCGAGCTCA	7020
Qy	7021	TCGAGGCCAACCTCTTGTGGCGCAGAGATGCGCGAAACATCACTCTGGGTGAGTCAAG	7080
Db	7021	TAGAGGCTAACCTCTCTGTGGAGCAGAGATGCGCGCAACATCAACAGGTTGAGTCAAG	7080
Qy	7081	AGAAATAGGTAGTAATTTCTGGACTCTTTGCAACCGCTTACGCGGAGGGGATGAGAGGG	7140
Db	7081	AGAAACAAAGTGGTGAATTTCTGGACTCTCTCGATCCGCTTGTGGCAGAGGAGATGAGCGGG	7140
Qy	7141	AGATATCCGTCGCGCGGAGATCCTCGCAAAATCCAGGAAGTTCCCTCAGCGTTCGCCCA	7200
Db	7141	AGGTCTCCGTAACCCGAGAAATTTCTCGGAAGTCTCGGAATTCGCCCGGCCCTGCGCG	7200
Qy	7201	TATGGGCACCGCGGACTCAATCTCTCACTGTCTAGAGTCTCTGGAAGAACCCGACTACG	7260
Db	7201	TTTGGGGCGCGGAGTCAACACCCCGCTAGTAGAGAGCTGGAAAGCCTGACTACG	7260
Qy	7261	TCCTTCGGTGTACACGGATGCCATTTGCCACTACCAAGGCTCTCCAATACCACCTC	7320
Db	7261	AACCACCTGTGGTCCATGTGCTGCCGCTACACCTCCACGGTCCCTCTGCTGCCG	7320
Qy	7321	CACGAGAAAGAGACGGTTGTCTGACAGAAATCAATGCTTCTGCGCTTGGCGGAGC	7380
Db	7321	CTCGAAAGAGCGTACGGTGTCTCAACCGAATCAACCTTACTACTGCTTGGCGGAGC	7380
Qy	7381	TCGCCATAAGACCTTCGGTAGCTCCGGATCGTCGGCGGTTGATAGCGGACGCGACCG	7440
Db	7381	TTGCCACCAAGTTTGGCAGCTCTCAACTTCCTCGGCATTTACGGGCGCAATACGCAA	7440
Qy	7441	CCCTTCTGACCTTGGCTCCGACAGCGTGACAAAGGATCCGAGCTTGAGTTCGTACTCCT	7500
Db	7441	CATCTCTGAGCGCGCCCTTCTGGCTGCCCGCCGACTCGAGCTTGAGTCTCTATCTT	7500
Qy	7501	CCATGCCCGCTTGAAGGGAGCGGGGGACCCGATCTCAGCAGCGGTTCTTGGTCTA	7560
Db	7501	CCATGCCCGCTTGGAGGGGAGCGCTGGGATCCGGATCTCAGCGACGGTCACTGTGCGA	7560
Qy	7561	CCGT---GAGTGAAGAGGCTAGTGAGGATGTGCTGCTGCTCAATGCTCTATACGTGA	7617
Db	7561	CGGTCACTAGTGGGGCGACACGGAAGATGTGCTGCTGCTCAATGCTCTATTCCTGA	7620
Qy	7618	CAGCGCCCTGATCACGCCATGCGCTCGGAGGAAGTAAGCTGCCCATCAACCCGTTGA	7677
Db	7621	CAGCGCACTCGTCAACCCCGTGGCTCGGGAAGAACAAATCTGCCATCAACGCACTGA	7680
Qy	7678	GCAACTCTTGTCTCGCTACCAACACATGCTTACGCCACAACATCCCGCAGCGCAAGCC	7737
Db	7681	GCAACTCTTGTCTAGCCATCAAACTCTGGTGTATTCACCACTTCAGCAGTCTTGGCC	7740
Qy	7738	TCCGCGAAGAGGTCACCTTTGACAGATTGCAAGTCTCTGGATGATCATTTACCGGACG	7797
Db	7741	AAAGCGAGAAAGTCACTTTGACAGATGCAAGTCTGCGACAGCCATTACAGGACG	7800
Qy	7798	TACTCAAGGAGATGAAGCGAAGCGTCCACAGTTAAGGCTTAAGCTTCTATAGAGG	7857
Db	7801	TGCTCAAGGAGTCAAAAGCAGCGGCTCAAAAGTGAAGGCTAACTTGCTATCCGCTAGAGG	7860
Qy	7858	AGGCTTGAAGCTGACGCCCGCCACATTCGGGCAATCCAAATTTGGCTATGGGGCAAGG	7917
Db	7861	AAGCTTGACGCTGACGCCCGCCACATTCAGGCAAAATCCAAATTTGGCTATGGGGCAAAAG	7920
Qy	7918	ACGTCCGGAACCTATCCAGAGCGCGTTAACCCATCCGCTCGCTGTTGGGAGACTTGC	7977
Db	7921	ACGTCCGTTGCCATGCGCAAGAGCGGTAGGCCACATCACTCCGTTGGGAAGACCTTC	7980
Qy	7978	TGGAAGACACTGAAACACCAATTTGACACCACTCATGGCAAAAGTGAAGTCTTCTGCG	8037

Db	7981	TGGAAGACAGTGTAAACCAATAGACACTACCATCATGCGCAAGAACGAGGTTTTCTGCG	8040
Qy	8038	TCCAAACGAGAGGAGCGCGCAGCTCGCTTATCGTATTTCCAGACACTGGAG	8097
Db	8041	TTCAGCCTTGAGAAAGGGGGTTCGTAAGCCAGCTCGTCTCATCGTGTTCGCCAGCTGGCG	8100
Qy	8098	TTTCGTGTATGCCGAGAGATGGCCCTTTACGAGTGTCTCCACCCTTCCTCAGCGCGTGA	8157
Db	8101	TGCGCGTGTGCGAGAAAGATGGCCCTGTACGACGTGTAGCAAGCTCCCCCTGGCGTGA	8160
Qy	8158	TGGGCTCTCATACGAGATTTCAATCTCCCAAGCAGCGGGTCGAGTTCCTTGGTGAATA	8217
Db	8161	TGGGAAGCTCTACGATTTCCATATCTACCAGGACAGCGGGTTGAATTCCTCGTGAAG	8220
Qy	8218	CTTGGAAATCAAGAAATGCCCTATGGGCTTCTATATGACACCCCGCTGTTTGTACTCAA	8277
Db	8221	CGTGAAGTCCAAAGAGACCCCGATGGGTTCTCGTATGATACCCGCTGTTTGACTCCA	8280
Qy	8278	CGTCACTGAGAGTGAATTCGTTTGAAGAGTCAATTTACCAATGTTGTGACTTGGCCC	8337
Db	8281	CAGTCACTGAGAGCGACATCCGTAACGAGAGGCAATTTACCAATGTTGTGACTGGACC	8340
Qy	8338	CCGAGCCGACAGAGCGCATAAAGTCTCTACAGAGCGCTTTACATCGGGGGTCCCTGA	8397
Db	8341	CCCAAGCCCGGTGGCCCATCAAGTCCCTCACTGAGAGGCTTTATGTTGGGGGCCCTCTTA	8400
Qy	8398	CTAACTCAAAAGGGAGAACTCGGTTATCGCGGTTCGCCGCAAGTGGCGTGTGACGA	8457
Db	8401	CCAAATCAAGGGGGGAAACTCGGCTACCGAGGTGCCGCGGAGGGGTACTGACAA	8460
Qy	8458	CTAGCTCGGTAATACCTTCAATGTTTACGAGGCACTGAGGCTGTGAGCTGCAAA	8517
Db	8461	CTAGCTGTGTAACACCTTCACTTGTACATCAAGGCGCGGCGAGCCGTGCGAGCGCAG	8520
Qy	8518	AGCTCAGAGCTGACGATGCTCGTGAAACGAGAGGACCTTGTGCTTATCTGTGAAAGCG	8577
Db	8521	GGCTCAGGAGCTGACCATGCTGCTGTGTCGCGACGACTTAGTCTGTTATCTGTGAAAGTG	8580
Qy	8578	CGGGAACCGAGGAGTGGCGGCGCTTACAGGCTTACGAGGCTATGACTAGGTATT	8637
Db	8581	CGGGGTTCAGGAGGACGCGCGAGCTTGAGAGCTTTCAGGAGGCTATGACCAAGTACT	8640
Qy	8638	CCGCCCCCGGGGATCCCGCCCAACAGAAATACGACTCGAGCTGATAACATCATGTT	8697
Db	8641	CCGCCCCCGGGGACCCCGCCACACAGAAATACGACTTGGAGCTTATAACATCATGCT	8700
Qy	8698	CTTCAATGTGTGAGTGTGCGCAGCATGCTCTGTGCAAAAGGTTATATCTCCTACCCGTG	8757
Db	8701	CCTTCAACGTTGTCAGTGTGCGCCACGAGCGCTGGAAGAGGGTCTACTACCTTACCCGTG	8760
Qy	8758	ACCCACCAACCCCTTTGCAACGGGCTGCTGGGAGACAGCTAGACACACTTCAATCAACT	8817
Db	8761	ACCTTCAACCCCTTCGCGAGAGCCGCTGGGAGACAGCAAGACACACTCCAGTCAATT	8820
Qy	8818	CTTGGCTTAGGCAATATCATGATGTCGCCACCTTATGGCAAGGATGATTTCTGTATGA	8877
Db	8821	CCTGGCTTAGGCAATATCATGATGTCGCCACCTTATGGCAAGGATGATGATGATGATGA	8880
Qy	8878	CTCACTTTTCTCCATCTTCTAGCTCAAGAGCAACTTGAAGAGCCCTGATTTGTGAGA	8937
Db	8881	CCCATTTCTTTAGGCTCTCATAGCCAGGATCAGCTTGAACAGGCTTTTAACTGTGAGA	8940
Qy	8938	TCTACGGGGTGTGCTTCTTCCATTTGAGCCACTTGAACCTTACCTCAGATCATTTGAACGACTCC	8997
Db	8941	TCTACGAGGCTGTCTTCTTCCATAGAACCACTGGATCTACCTCCAATCATTTCAAAGACTCC	9000
Qy	8998	ATGCTTTAGCGCAATTTTACTTCCACAGTCTCTCCAGGTTGATCAATAGGGTGGCTT	9057
Db	9001	ATGSCCTCAGCGCATTTTCTCTCCACAGTTACTCTCCAGGTTGAATCAATAGGGTGGCG	9060
Qy	9058	CATGCTCAGGAACTTTGGGGTACACCTTTGCCAACTTGAGACATCGGGCCAGAGTG	9117
Db	9061	CATGCTCAGAAACTTTGGGGTCCCGCTTTCGAGCTTGGAGACACCGGGCCCGGAGCG	9120

Db 1141 GGAGCCGCACTCTCTCGGCCCTCTAGTGGGGACCTGTGCGGGTCTGTCTTCTTGG 1200
 Qy 1201 TCTCCAGCTGTTCACCTTCTCGCTCGCCGCGATGAGACAGTGCAGGACTGCAACTGCT 1260
 Db 1201 TTGGTCAACTGTGTACCTTCTCTCCAGCGGCCACTGGACGACGCAAGACTGCAATTGT 1260
 Qy 1261 CAATCTATCCCGGCATGTATCAGGTACCGGATGGCTTGGGATATGATGATGAAGTGT 1320
 Db 1261 CTATCTATCCCGGCCATATAAAGGGTTCATCGCATGGCATGGGATATGATGATGAAGTGT 1320
 Qy 1321 CACCTACAAAGCCCTAGTGGTGTGCGAGTGTCTCGGATCCCAAGCTGTGCTGGACA 1380
 Db 1321 CCCCTACGCGAGGTTGGTGGTAGCTCAGCTGTCTCGGATCCCAAGCCATCATGGACA 1380
 Qy 1381 TGGTGGCGGGGCCCATCTGGGAGTCTTGGCGGCTTGGCTACTATTCATGGTAGGGA 1440
 Db 1381 TGATCGCTGGTGTCTACTGGGAGTCTTGGCGGCATAGCGTATTTCTCCATGGTGGGA 1440
 Qy 1441 ACTGGCTAAGGTTCTGATTGTGGCGTACTCTTTGCGGCGTTGAGCGGGAGACCCACA 1500
 Db 1441 ACTGGCGAAGTCTCTGGTAGTCTGTCTGTCTATTTGCGGCGTCTGACGCGGAAACCCACG 1500
 Qy 1501 CGACGGGAGGTTGGCGGCCACACCACTCTCGGCTTTCAGTCCCTTTCTCATCTGGGG 1560
 Db 1501 TCACCGGGGAAGTGGCGGCCACACAGCGCTGGGCTTGTGTCTCTTACACAGGCG 1560
 Qy 1561 CGTCTCAGAAATCCAGCTTGTGAATACCAACGCGAGCTGGGCATCAACAGGACTGCC 1620
 Db 1561 CCAAGCAGAAATCCAACTGTATCAACACCAACGCGAGTTGGGCATCAATAGCACGSCCT 1620
 Qy 1621 TAAATTGCAATGACTCCTCCAACTGGGTTCTTTGCGGCGCTGTTTACGCAACAACT 1680
 Db 1621 TGAATGCAATGAAAGCTTTAAACCGGCTGGTTAGCAGGCTCTTCTATCAGCACAAAT 1680
 Qy 1681 TCAACTCGTCCGGTCCCGAGCGCATGCGAGCTGCGCGCCATTTGCTGTTGGCCC 1740
 Db 1681 TCAACTCTTCAGGCTGTCTGAGAGGTTGGCCAGCTGCCAGCGCTTACCGATTTGGCCC 1740
 Qy 1741 AGGGTGGGGCCCATCACCTATACCTAAAGCTTAACAGCTCGGATCAGAGGCTTATTGCT 1800
 Db 1741 AGGGTGGGGTCTATCAGTTATGCCAACGGAAGCGCTCGACGACGCGCCCTACTGCT 1800
 Qy 1801 GGCAATACGCGCTCGACCTGTGGTGTGTGACCCCGGTGCGAGGTTGTGGTCCAGTGT 1860
 Db 1801 GGCACCTACCTCCAAAGCTTGTGGCATTTGCGCGCAAGAGCGTGTGTGGCCGGTAT 1860
 Qy 1861 ATTGTTTACCCCAAGCCCTTGTGTGGTGGGACCAACCGATCGTTCGGTGTCCCTACGT 1920
 Db 1861 ATTGCTTCACTCCAGCCCGTGGTGGGAAACGACCGACAGGTGGGGCGCGCTACCT 1920
 Qy 1921 ATAGCTGGGGGAATGAGACAGCTGATGCTCTCAACACACGCGTCCGCCACAAG 1980
 Db 1921 ACAGCTGGGTGCAATGATACGATGTCTTCGTCCTTAACACACGAGGCCCGCTGG 1980
 Qy 1981 GCAACTGTTTGGCTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Db 1981 GCAATGTTTGGTGTGATCTGGATGAATCAACTGATGATGATGATGATGATGATGATGAT 2040
 Qy 2041 CCGCGTGTAAATCGGGGGTGGTGAACCGACCTTGTATGCTGCGCCACGAGCTGCTTCC 2100
 Db 2041 CCGCGTGTATCGGAGGGTGGGCAACACCTTGTCTGTGCGCCCACTGATTTGCTTCC 2100
 Qy 2101 GGAAGCACCCGAGCTTACACAAATGTGCTGGGCGCTGGTGTGACACTAGGT 2160
 Db 2101 GCAAACTCCGGAAGCACATACCTGCTGGGCGCTCGGTCCTTGGATTAACCCAGGT 2160
 Qy 2161 GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTCCATCT 2220
 Db 2161 GCATGGTCGACTACCCGATAGGCTTTGGCACTATCTTGTACCATCAATTAACCATAT 2220
 Qy 2221 TTAAGTTAGGATGATGTGGGGGGCGTGGAGCAAGGCTCAATGCGCGCATGCAATTTGGA 2280
 Db 2221 TCAAGTCAGGATGATCGTGGGAGGGTTCGAGCAGAGCTGGGAGCGGCTCTGCGCGTAGGG 2280

Qy 2281 CTGAGGAGAGCGCTGTAACTTTGGAGGACAGGGATAGGTCAAGACTCAGCGCGCTGCTGC 2340
 Db 2281 CCGGGGGCAACGCTGTGTATCTGGAAGACAGGGACAGGTCCGAGCTCAGGCCCATTTGCTGC 2340
 Qy 2341 TGTCTACAAACAGAGTGGCAGATACCTCCCTGTGCTTTTACCACCTTACCGGCTTTATCCA 2400
 Db 2341 TGTCCACCACACAGTGGCAGGTCTTCCGTGTCTTTCACGACCTTGCACGCTTGTGCCA 2400
 Qy 2401 CTGGTTTGATTCATCTCCATCAGAACATCGTGGACGTGCAATACCTGTAGCGTGTAGGCT 2460
 Db 2401 CCGGCTCATCCACCTCCACCAACATTTGTGACGTGACGTACTTTGTACGGGTGGGGT 2460
 Qy 2461 CAGCGTTTGTCTCTTTGCAATCAATGGAGTACATCTGTGTGCTTTCTCTCTCTCTG 2520
 Db 2461 CAAGCATCGCTCTGGGCCATTAAGTGGAGTACGTCTCTCTCTCTCTCTCTCTCTG 2520
 Qy 2521 CAGACGCGCGTGTGCTGCTGCTGATGATGCTGTCTCTCTCTCTCTCTCTCTCTCTCTG 2580
 Db 2521 CAGACGCGCGTGTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2580
 Qy 2581 CTTAGAGAACTTGGTGTCTCAATGCGCGCTCTGCTGGCGGAGCGCATGTTCTCTCT 2640
 Db 2581 CTTTGGAGAACTCTGTAATACTCAATGACAGCATCTCTGCGCGGAGCGCATGTTCTGT 2640
 Qy 2641 CTTTCTTGTGTTCTTCTGCGCGCTGTGTACATTAAGGCGAGCTTGGCTCTCTGGGCGG 2700
 Db 2641 CTTCTCTGTGTTCTTCTGCTTTGCGTGTATCTGAAGGGTATGTTGGGTCTCTGGAGCGG 2700
 Qy 2701 CGTATCTTTTATGCGTATGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 Db 2701 TCTACGCTTCTACGAGGATGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 Qy 2761 CTTTACGCTTGGACCGGAGATGCTGCACTGTGCGGGGTGCGGTTCTTGTAGTCTGG 2820
 Db 2761 CATACGACTTGGACACGAGGTGCGCGTGTGTGCGGCGTGTCTTGTGCGGTTAA 2820
 Qy 2821 TATTCTTGACTTGTACCATACTACAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2880
 Db 2821 TGGCGCTGACTGTCTCACCATATTACAAGCGCTATATCAGCTGTGTCATGTGTGGTCTC 2880
 Qy 2881 AATCTTTATCACAGCGCGGCGCACATGCAAGTGTGGGTCTCTCTCTCTCTCTCTCTCTCT 2940
 Db 2881 AGTATTTTCTGACAGAGTAGAAGCGCACTGACGCTGTGGGTCTCTCTCTCTCTCTCTCT 2940
 Qy 2941 GGGAGCGCCGATGCGCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3000
 Db 2941 GGGGGGGCGGATGCGCTCATCTTACTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
 Qy 3001 ACATACCAAACCT 3060
 Db 3001 ACATACCAAACCT 3060
 Qy 3061 CGAGAGTGGCTTCT 3120
 Db 3061 TTAAGTCCCT 3120
 Qy 3121 AAGTGGCGGGGTCTATTATGTCCTTCAAAATGGTCTTCTATGAAGCTGGGCGCTGACAGTA 3180
 Db 3121 AGATAGCGGAGGTCTATTAAGTGCNAATGGCCATCATCAAGTGGGGCGCTTACTTGCA 3180
 Qy 3181 CGTACGTTTAAACCATCTTTACCCACTGGGGGACTGGGCCACGCGGCGCTTACGAGACC 3240
 Db 3181 CCTATGTATACCACTCTCAACCTCTCTCGAGACTGGGCGCACACGCGCTCTGCGAGATC 3240
 Qy 3241 TTGGGTGGCGGTAGAGCCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3300
 Db 3241 TGGCGGTGGTGTGGAAACCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3300
 Qy 3301 GAGCAGACACCGCTGTGGGGGACATCACTTTGGGTCTACCCGTCTCTCTCTCTCTCTCTCT 3360
 Db 3301 GGGCAGATACCGCGCGGTGGGTGACATCATCAACGCGCTTGGCGCTCTCTCTCTCTCTCT 3360

Db 5521 AGTTAAGCAGAAAGGCCCTCGGCTCCTGCGAGCCGCTCCGCCCAAGCAGAGGTTATCA 5580
Qy 5581 CTCCTGGTGGAGTCAAGTGGCGAGCCCTTGAGACCTTCTGGGGAAGCACAATGTGGA 5640
Db 5581 CCCCTGCTGCAGACCAACTGCGAGAACTCGAGGTCTTCTGGCGAAGCACAATGTGGA 5640
Qy 5641 ATTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCTGCTGGAACCCCGCA 5700
Db 5641 ATTTCATCAGTGGGATACATATTTGGCGGCGCTGTCAACGCTGCTGTGTAACCCGCA 5700
Qy 5701 TAGCATCATGATGGCATTTTACAGCTTCTATCACTAGCCGCTCACACCACAAAACACCC 5760
Db 5701 TTGCTTCATGATGGCTTTTACAGCTGCGCTACAGCCCACTAAACCTGCGCAACCC 5760
Qy 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGCTGCGT 5820
Db 5761 TCCTTTCAACATATTGGGGGGTGGGTGGCTGCCAGCTCGCGCCCGCGTGGCGCTA 5820
Qy 5821 CAGCTTTCTGGGCGCGGATCGCGGAGCGGTGTTGGCAGCATAGGCTTTGGGAAG 5880
Db 5821 CGCCTTTGTGGCGCTGCTTAGCTGGCGCGCATCGGCAGCGTTGGACTGGGAAG 5880
Qy 5881 TGCTGTGGACATCTTGGCGGCTATGGGCGAGGGTAGCGGCGCACTCGTGGCTTTA 5940
Db 5881 TCCTGTGGACATCTTGGCAGGGTATGGCGGGCGTGGCGGAGCTCTTGTAGCATCA 5940
Qy 5941 AGGTATGAGCGGCGGCTCCACCGAGGACCTGGTCAACTTACTCCCTGCCATCC 6000
Db 5941 AGATCATGAGGCTGAGTCCCTCCACGAGGAGACCTGGTCACTGCTGCCGCTATC 6000
Qy 6001 TCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6060
Db 6001 TCTGCTGAGCCCTTGTAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6060
Qy 6061 GCGCGGAGAGGGGCTGTGCACTGATGATGATGATGATGATGATGATGATGATGATG 6120
Db 6061 GCGCGGCGAGGGGCGAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 6120
Qy 6121 ACCAGCTCCCTTACGACCTATGCTGCTGAGAGCGAGCTGACGACGCTGCTACTCAGA 6180
Db 6121 ACCATGTTTCCCCCAGCGACTATGCTGCGGAGAGCGATGACGCGCGCGCTACTGCCA 6180
Qy 6181 TCCTCTCTAGCCTTACCATCACTCAACTGCTGTAAGCGGCTCCACAGTGGATTATGAGG 6240
Db 6181 TACTCAGCAGCTCACTGTAAACCCAGCTCTCTGAGCGAGCTACATCAGTGGATAGCTCG 6240
Qy 6241 ACTGCTCTAGCCATGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6300
Db 6241 AGTGTACCACTCCATGCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6300
Qy 6301 TGTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCCTCGCGGTTACCGGGAGTCC 6360
Db 6301 TGCTGAGGACTTTAAGACCTGGCTGAAAGCAAGCTCATGCCAACCTGCTGGGATTC 6360
Qy 6361 CTTTCTCTCATGCCAAGCGGCTACAGGAGTCTGGCGGGGAGCGGCACTATGCAAA 6420
Db 6361 CTTTGTGCTCTGCCAGCGCGGTATAGGGGCTTGGCGGAGAGAGCGCAATATGCACA 6420
Qy 6421 CCACCTGCCATTCGGAGCAGATGCGCGGACATGTCAAAAAGGTTTCCATGAGGATCG 6480
Db 6421 CTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAAGGACGATGAGGATCG 6480
Qy 6481 TAGGCGCTAGAACCTGCGACACACGTTGGCAGGAACTTCCCCATCAACCATACACA 6540
Db 6481 TCGGTCTTAGAACCTGCGAGAACATGTGGAGTGGGACGTTCCCCATTAACCGCTACA 6540
Qy 6541 CGGACCTTGCACACCTCCCGCGCGCCAACTATTCCAGGCGCTATGCGGGGTGGCTG 6600
Db 6541 CGGCGCTTACTCTCTCTGCGCGAACATTAAGTTGCGCTGTGGAGGCTGTCTG 6600
Qy 6601 CTGAGGAGTACGTGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGCGCATGACA 6660
Db 6601 CAGAGGAATACGTGGAGATAGCGGGTGGGGACTTCCACTACGTATCGGGTATGACTA 6660

Qy 6661 CTGACAAAGTATGCTGCCATGCGAGGTTCCGCCCCCGAAATTTCTTCACGGAGGTGATG 6720
Db 6661 CTGACAAATTTAAATGCCCCGTCAGATCCCATGCCCCGAAATTTTTCACAGAATTTGAGC 6720
Qy 6721 GAGTGGGTTCACAGGTACGCTCCGCGCTGCAAACTCTTCTTACCGGAGGAGCTCACGT 6780
Db 6721 GGTGGCGCTCATAGGTTTGGCCCCCTTTCGCTGCGGAGGAGGTATCAT 6780
Qy 6781 TCAGGTTCGGCTCAACCAATATCTTGGTTCGGGTGCGAGCTCCCATCGGAGCCGAAACCG 6840
Db 6781 TCAGAGTAGGACTCCACGAGTACCCGCTGGGTGCGAATTTACCTTGGAGCCGAAACCG 6840
Qy 6841 ACCTAACAGTGTACTTCCATGCTCACCGATCCCTCCACATTAACAGCAGACACGCTA 6900
Db 6841 ACCTAGCCGTGTGAGCTCCATGCTCCTGATCCCTCCCATATAACAGCAGAGCGGCG 6900
Qy 6901 AGCTAGGCTGGCTAGAGGTCTCCCTCTTCTTACGAGCTCATCAGCTAGGCTAGTTGT 6960
Db 6901 GGAGAGGTTGGCAGAGGCTCACCTCTTCTATGCGCAGCTCTCGGCCAGCAGCTGT 6960
Qy 6961 CTGCGCTCTTTTGAAGGCGACATGCACTACCCATGATCCCTCCCGGAGCTGACTCA 7020
Db 6961 CGCTCCATCTCTCAAGGCAACTTGCACCGCCAACTGACTCCCTGACGCGAGCTCA 7020
Qy 7021 TCAGGCGCAACCTCTTGTGGCGCAGAGATGGCGGAACATCACTCGCTGAGCTAG 7080
Db 7021 TAGAGCTTAACCTCTTGTGAGGCGAGAGATGGCGGCAACATCACAGGCTTGTAGTCA 7080
Qy 7081 AGAATAGGTAGTAATTTCTGAGCTCTTTCGAAACGCTTTCACGCGAGGGGATGAGAGG 7140
Db 7081 AGAACAAGTGGTGTATTTCTGAGCTCTTTCGATCGCTTGTGGCAGAGGAGATGAGCGG 7140
Qy 7141 AGATATCCGTTCGCGCGGAGATCTCGGAAATTCAGGAAATTCCTCAGCTGAGCTGCTCA 7200
Db 7141 AGCTCTCCGTACCCGAGAAATTTCTCGGAAATTCGAGATTCGCGCCAGCTGCTGCGC 7200
Qy 7201 TATGGGCAACCCCGAGCTCAATCTCTCCTAGCTAGAGTCTTGGAGGAGACCCGAGCTAG 7260
Db 7201 TCTGGGCGCGCGGAGCTCAACCCCGCTAGTAGAGCTGGAAGAGCTGGAAGAGCTGACTAG 7260
Qy 7261 TCCTTCGGTGTACAGGATGCCATTCACCTCAAGGCTCTCAATACCACTC 7320
Db 7261 AACCACTGTGGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7320
Qy 7321 CACGAGAAAGAGAGCGGTTGCTCTGACAGAAATCCAAATGTCTTCTGCTGCTGCTGCTG 7380
Db 7321 CTGGAAGAGAGCTACGCTGCTCTCACCGAATCAACCTATCTACTGCTTGGCCGAGC 7380
Qy 7381 TCGCACTAAGACTTCGCTGAGCTTCGAGTCTGCTGCGCTGTTGATAGCGGCGAGCGAGC 7440
Db 7381 TTGCCACCAAAAGTTTGGCAGCTCTCAACTTCGGCATTTACGGGCGCAATACGACAA 7440
Qy 7441 CCCTTCTGACTGCTGCTCCGACGACGAGTGAAGGATCCGAGCTTGTAGTCTGCTCT 7500
Db 7441 CATCTCTGAGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7500
Qy 7501 CCATGCCCCCTTGAAGGAGCGCGGAGACCCGATCTCAGCGAGCGGCTTCTGCTGCTA 7560
Db 7501 CCATGCCCCCTTGAAGGAGCGCGGAGCTGCGGATCTCGGATCTCAGCGAGGCTCATG 7560
Qy 7561 CCGT - - - GAGTGAAGGCTAGTGAAGATGCTGCTGCTCAATGCTCTATAGTGA 7617
Db 7561 CCGT CAGTGTGGGCGCGACAGGAGATGCTGCTGCTGCTCAATGCTCTATTCCTGA 7620
Qy 7618 CAGCGCGCTGATCAGCGCATGCTGCGGAGGAAAGTAACTGCCATCAACCGTGA 7677
Db 7621 CAGCGCACTCTGCTCAGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7680
Qy 7678 GCAACTCTTCTGCTGCTCAGCAACATGCTTACGCCACAACTCCCGCAGCGCAAGCC 7737
Db 7681 GCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7740

ORGANISM: Hepatitis C virus
US-10-475-989-2

Query Match 65.2%; Score 6258; DB 7; Length 9622;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 7545; Conservative 0; Mismatches 2050; Indels 27; Gaps 3;

QY 1 GCCAGCCCTGTATGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAACACTG 60
DB 1 GCCAGCCCTGTATGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAACACTG 60
QY 61 TCTTACGCGAAGAGCTTAGCCATGGCGTGTAGTAGTATGAGTGTGTCAGCCTCCAGGAC 120
DB 61 TCTTACGCGAAGAGCTTAGCCATGGCGTGTAGTAGTATGAGTGTGTCAGCCTCCAGGAC 120
QY 121 CCCCCCTCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTGCCAG 180
DB 121 CCCCCCTCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTGCCAG 180
QY 181 GAGCAGCGGTCTCTTCTTGATCAACCCGCTCAATGCTCGAGATTGGCGGTGCCCC 240
DB 181 GAGCAGCGGTCTCTTCTTGATCAACCCGCTCAATGCTCGAGATTGGCGGTGCCCC 240
QY 241 GCGAGATGCTAGCCGAGTAGTGTGGGTTCGGAAGCCCTTGTGTACTGCTGTAGG 300
DB 241 GCAAGACTGTAGCCGAGTAGTGTGGGTTCGGAAGCCCTTGTGTACTGCTGTAGG 300
QY 301 GTGCTTCGGAGTCCCGGGAGGTCTGTAAGCGTGCACCATGAGCAGCAATCCTAAAC 360
DB 301 GTGCTTCGGAGTCCCGGGAGGTCTGTAAGCGTGCACCATGAGCAGCAATCCTAAAC 360
QY 361 CTAAAGAAACCAACAGCTAAACCAACCGCCGCCACAGGAGTCAAGTTCCCGGGCG 420
DB 361 CTAAAGAAACCAACAGCTAAACCAACCGTCCGCCACAGGAGTCAAGTTCCCGGGCG 420
QY 421 GTGCTCAGATGTTGTGGAGTTTACTGTTTGGCGGAGGGGCCACAGTTGGGTGTGC 480
DB 421 GCGGTACAGTCTGTTGTGGAGTTTACTGTTTGGCGGAGGGGCCCTAGATTGGGTGTGC 480
QY 481 GCGGACTAGGAAGCTTCGAGCGGTGCAACCTCGTGAAGGCGCAACCTATCCAA 540
DB 481 GCGGAGAGGAAGCTTCGAGCGGTGCAACCTCGTGTGAGACGTGAGCTATCCCA 540
QY 541 AGGCTCGCGACCGAGGCGAGGCTGGGTCTCAGCCCGGTACCTTTGGCCCTCTATG 600
DB 541 AGGCACTGCGCCGAGGCGAGGCTGGGTCTCAGCCCGGTACCTTTGGCCCTCTATG 600
QY 601 GCAATGAGGCTTGGGTGGCAGGATGGTCTCTGTCAACCGCGGTCTCCGGCTAGTT 660
DB 601 GCAATGAGGCTTGGGTGGCAGGATGGTCTCTGTCTCTCCCGGTCTCCGGCTAGCT 660
QY 661 GGGGCCCCACGAGCCCCGGGTAGGTGCGTAACTTGGGTAAAGTCAATGATACCCCTTA 720
DB 661 GGGGCCCCACAGCCCCGGGTAGGTGCGCAATTTGGGTAAAGTCAATGATACCCCTTA 720
QY 721 CATGGGCTTCGCGATCTCATGGGTACATTCGCTCGTGGCGCCCTAGGGGCG 780
DB 721 GTGGCGCTTTCGCGACCTCATGGGTACATACCGCTCGTGGCGCCCTCTTGGAGCG 780
QY 781 CTGCAGGCGCTTGGCACAACGGTGTCCGGTTCGTGAGGACGGCGTGAATATGCAACAG 840
DB 781 CTGCAGGCGCTTGGCAGCATGGCGTCCGGTTCGTGAGGACGGCGTGAATATGCAACAG 840
QY 841 GGAACCTTCCGGTGTCTTCTCTATCTCTCTGCTCTGCTCTGCTCTGCTCTGCTGACCA 900
DB 841 GGAACCTTCCGGTGTCTTCTCTATCTCTCTGCTCTGCTCTGCTCTGCTCTGCTGACCA 900
QY 901 TCCAGCTTCGCTTATGAGTGCAGGCTCCGGATATACCATGTACGAGCACT 960
DB 901 TGCCGCTTACGCTTACCAAGTGCAGATTCCTCGGGGCTTACCATGTACCAATGATT 960
QY 961 GCTCCTCAACTGAGCAATTGTATAGGAGCGGAGCTGATCATGCTACTCCCGGGTGC 1020
DB 961 GCTCCTCAACTGAGCAATTGTATAGGAGCGGAGCTGATCATGCTACTCCCGGGTGC 1020

DB 961 GGCCTAATTGAGTATTGTGTACAGGCGCGATGCCATCTCTGCACACTCCGGGTGTG 1020
QY 1021 TCCCTGTGTTTCCAGGAGGTAAACAGCTCCGTTGTGGGTAGCGCTCACTCCACACCTCG 1080
DB 1021 TCCCTGTGTTTCCAGGAGGTAAACAGCTCCGTTGTGGGTAGCGCTCACTCCACACCTCG 1080
QY 1081 CGGCCAGGAATGCCAGGTCCTCCACTACGACATAGACGCGACGTCGACTCTCGCTGTTG 1140
DB 1081 CCACCAGGACGGCAAACTCCCAACAGCAGCTTCGACGTATATCGATCTGCTGTTGTCG 1140
QY 1141 GAGCGCTGCTTCTCTCGCTATGTAGCTGGGGATCTCTCGGGATCTATTTTCTCTCG 1200
DB 1141 GAGCGCTGCTTCTCTCGCTATGTAGCTGGGGATCTCTCGGGATCTATTTTCTCTCG 1200
QY 1201 TCTCCAGCTGTTCACCTTCTCGCTCCGCGCATGAGACAGTGCAGGACTGCAACTGCT 1260
DB 1201 TCGTCAACTGTTTACCTTCTCTCCAGGCGCATCTGCAGCAGCAAAAGTGCAAATGTT 1260
QY 1261 CAATCTATCCCGCATGTATCAGGTCAAGCATGCTGGGATGGGATATGATGATCAACTGCT 1320
DB 1261 CTATCTATCCCGCATATTAACGGGTCACTGCATGGCATGGGATATGATGATCAACTGCT 1320
QY 1321 CACTTCAACAGCCCTAGTGTGTCTCGAGTTGCTCCGGATCCCAAGCTGTCTGTGACA 1380
DB 1321 CACTTCAACAGCCCTAGTGTGTCTCGAGTTGCTCCGGATCCCAAGCTGTCTGTGACA 1380
QY 1381 TGTGCGGGGGCCACTGCGGAGTCTTGGCGGCTTGTGCTACTATTTCCATGTTAGGA 1440
DB 1381 TGTGCGGGGGCCACTGCGGAGTCTTGGCGGCTTGTGCTACTATTTCCATGTTAGGA 1440
QY 1441 ACTGGGCTAAGTTTCTGATGTGGCGTACTCTTTCGGGGTGTGAGGGAGACCCACA 1500
DB 1441 ACTGGGCGAAGGCTCTGTGTAGTCTGCTGCTATTTGCGCGGCTGACGCGGAAACCCACG 1500
QY 1501 CGACGGGAGGGTGGCGGCCACACACTCTCCGGTTTCACTCTCTCTTCTCTCTCTG 1560
DB 1501 TCACGGGGGAAGTGGCGGCCACACAGGCTGGGCTTGTGCTCTCTTACACGAGCG 1560
QY 1561 CGTCTCAAGAAATCCAGCTTGTGAATACCAACGGCAGCTGGCAGATCAACAGGACTGCC 1620
DB 1561 CCAAGCAGAACATCCAACTGATCAACCAACGGCAGTGGCAGATCAATAGCAGCGCT 1620
QY 1621 TAAATTGCAATGACTCCCTCAAACTGGGTCTTTCGGCGCTGTTTTAGCGACACAAGT 1680
DB 1621 TGAATCTCAACAGTAGCTTTACCAACGGCTGGTTAGCAGGGCTCTTCTATCGCCACAAT 1680
QY 1681 TCAACTCGTCCGGTGGCGGAGCGCATGCGAGCTGCGCGCCCATTTGACTGGTTCGCC 1740
DB 1681 TCAACTCTTTCAGGCTGTCTTACCAACGGCTGGTTAGCAGGGCTCTTCTATCGCCACAAT 1740
QY 1741 AGGGTGGGGCCCCCATCACTATATAAGCTTAAAGCTTAAAGCTCAGAGGCTTATTTGCT 1800
DB 1741 AGGGTGGGGTCCCATCAGTTATGCAACGGAGCGGCTTACAGCAACGCGCTTACTGTT 1800
QY 1801 GGCATTAACGCGCTTCGACCGTGTGGTGTGTACCCGCTGCGAGGTGTGGTGTGCTCAGTGT 1860
DB 1801 GGCATTAACCTTCCAAAGCTTGTGGCATTTGTGCCCGCAAGAGCGTGTGTGGCGCGGTAT 1860
QY 1861 ATTGTTTACCCCAAGCCCTTGTGTGGTGGGACCAACGATTCGTTTCCGTTCTCCTTACGT 1920
DB 1861 ATTGTTTACCTCCAGCCCGCTGTGTGGTGGGAAACGACGACGATTCGGGCGGCTTACT 1920
QY 1921 ATAGCTGGGGGAGATGAGACAGCTGATGCTCTCAACCAACGCGCTCCGCCACAAG 1980
DB 1921 ACAGCTGGGTGCAATGATACGATGTCTTCGTTCTTAAACACAGGCGCACCGCTGG 1980
QY 1981 GCAACTGCTTCCGCTGTACATGGAATAGTACTGGTTCACTAAGACGTGCGGAGGTC 2040
DB 1981 GCAATGCTTCCGTTGCTTACCTTGGATGAATCACTGATGATTCACCAAGTGTGCGAGCGC 2040
QY 2041 CCGCGTGTACATCGGGGGGTCCGTAAACCGCACTTGTATCTGCCCCACGAGCTCTTCC 2100
DB 2041 CCGCTTGTCTACGAGGGGTGGGCAACAACTTGTCTCTGCCCACTGATTCCTTCC 2100

4261	Db		GAGGCGCTTATGACATAATAATTTGTGACGAGTGGCCATCTCCACGGATGCCATCCATCT	4320
4321	Qy		TGGGCATCGGCACAGCTCTGGACCAAGCGGACGCGCTGGAGCGGGCTCGTCTGCTCG	4380
4321	Db		TGGGCATCGGCATCTGCTTGTGACCAAGACGAGACTGCGGGGGCGAGATTGGTTGTGCTCG	4380
4381	Qy		CCACCGCTACACCTCCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAAATAGGCC	4440
4381	Db		CCACTGCTACCCCTCCGGGCTCGTCACTGTGTCCCATCTTAACATCGAGGAGGTTGCTC	4440
4441	Qy		TGTCCAAATATGAGAGATCCCTTCTATGTGCMAAGCCATCCCATTTGAGGCCATCAAGG	4500
4441	Db		TGTCCAACACCGGAGAGATCCCTTCTTACGGCAAGGCTATCCCTCTCGAGGTGATCAAGG	4500
4501	Qy		GGGGGAGGCATCTCATTTCTGTCATTTCCAAGAAAGAAATGTGACAGAGCTCGCGCAAGC	4560
4501	Db		GGGGAAGACATCTCATCTTCTGTCACTCAAGAAAGAAAGTGCAGACGAGCTCGCGGAGC	4560
4561	Qy		TGACAGGCTTCGGACTGAACGCTGTAGCATATTACCGGGGCGCTTGATGTGTCGTCATAC	4620
4561	Db		TGGTCGCAITGGGCATCAATGCCGTGGCCTACTTACC CGGACTTGACGTGTCTGTCATCC	4620
4621	Qy		CGCCTATCGGAGACGTGGTTGTTCGTGCGCAACAGACGCTCTAATGACGGGTTTCA	4680
4621	Db		CGACAGCGCGATGTGTGCTGTGTGCAACCGATGCTCTCATGACTGGGCTTTACCGGGC	4680
4681	Qy		ATTTTGACTCAGTGATCGACTGCAATACATGCTCACCCAGACAGTGCAGCTTCAGCTTCG	4740
4681	Db		ACTTCGACTCTGTGATAGACTGCAACACGTGTGTCACTCAGACGTCGATTTTCAGCCTTG	4740
4741	Qy		ATCCACCTTTCACCATTTGAGACGACGACCGTGGCCCAAGACGCGGTGTGCGCTCGCAAC	4800
4741	Db		ACCCTACTTTACCATTTGAGACCAACGAGCTCCCCCAGGATGCTGTCTCCAGGACTCAGC	4800
4801	Qy		GGGAGGTGAACTGTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTTCGAGGAGAAC	4860
4801	Db		GCCGGGCGAGACTTGGCAGGGGGAAGCAGGCATCTACAGATTTGTGGCACCGGGGGAGC	4860
4861	Qy		GGCCCTCGGCATGTTCCGATCTTTCGGTCTGTGTGAGTGCATGACGGGGCTGTGCTT	4920
4861	Db		GCCCTTCGGGATGTTTCGACTCGTCTGTGTGAGTGTGTGAGTGTGTGAGTGTGTGCTT	4920
4921	Qy		GGTATGAGCTCACCGCGCTGAGACCTCGGTGTAGGTTTTCGGGCTTACTATAATACACCAAG	4980
4921	Db		GGTATGAGCTCATGCGCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGG	4980
4981	Qy		GGTTGCCGCTGTCAGGACCATCTGAGTTCTGGGAGAGCGCTTTCACAGGCGCTCACCC	5040
4981	Db		GGCTTCCGCTGTGCAGGACCATCTTGAAATTTTGGGAGGGCGTCTTTACGGGCGCTCACCC	5040
5041	Qy		ACATAGATGCCACTTCTCTGTCAGACTAAACGAGCAGGAGACAATTTTCCTTACTCTGG	5100
5041	Db		ATATAGATGCCACTTTCTATCCAGACAAGACGAGTGGGGGAACTTTTCCTTACTCTGG	5100
5101	Qy		TGGCATATCAAGCTACAGTGTGCCGAGGCGCTCAAGCTCCACCTCCATCTGTGGGACCAAA	5160
5101	Db		TAGCGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGGCCCTTCCCCCATCTGTGGGACGAG	5160
5161	Qy		TGTGGAAGTGTCTCATCGGCTGAAACCTACACTGCAAGGGGCCAACACCCCTCTGTGATA	5220
5161	Db		TGTGGAAGTGTGATCGCGCTTAAACCCACCCCTCCATGGGCCAACACCCCTCTGTATACA	5220
5221	Qy		GGCTAGGAGCGGTCCAAATATGAGTCTATCTCACACACCCCATTAATATCATCATG	5280
5221	Db		GACTGGGGCGCTGTTTCAGAAATGAAGTACCCCTGAGCGCAACCAATCACCAATAATCATATGA	5280
5281	Qy		CATGCACTGCGGCTGACCTGAGGTGCTGCTAGCACCTGGGTGCTGTGGTAGGCGGAGTCC	5340
5281	Db		CATGCACTGCGGCGACCTGGAGTGTGTACAGACACTGGGTGCTGTGGGGCGGTCC	5340
5341	Qy		TTGCAGCTTTGGCGCGACTATGCTCTGACGACAGGACAGTGTGGTCAATTTGTGGGACGATCA	5400

5341	Db	TGCGTCTGTGGCCGCGTATTTGCTCTCAACAGGCTGCGTGTCTCATAGTGGCAGGATTG	5400
5401	Qy	TCCTTGTCCGGGAAGCCAGCTGTCTGTTCCCGACAGGGAAGTCTCTTACACAGGAGTTTCGATG	5460
5401	Db	TCCTTGTCCGGGAAGCCGCAATTATACCTTGACAGGGAGGTTCTTACACAGGAGTTTCGATG	5460
5461	Qy	AGATGGAAGAGTGTGCTCACAACCTTCTTATCANTCGACGAGGAATGCACTGCGCGAGC	5520
5461	Db	AGATGGAAGAGTGTCTCTCAGCACTTACCGGTACATCGAGCAAGGATGATGCTCTGCTGAGC	5520
5521	Qy	AATTTCAAGCAAAAGGCGCTCGGGTGTGTCAAACGGCCACCAAGCAAGCGGAGCGCTGCTG	5580
5521	Db	AGTTCAAGCAGAAGCGCTCGGCTCTCGAGACCGGCTCCCGCCATGCAGAGTTTATCA	5580
5581	Qy	CTCCGCTGTGTGAGTCCAAGTGGCGAGCCCTTGAGACCTTCTGGGGGAAGCACATGTGA	5640
5581	Db	CCCGTGTCTCCAGACCAACTGGCAGAACTCGAGGTCTTCTGGGGGAAGCACATGTGA	5640
5641	Qy	ATTTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCTCGCTGGAACCCCGCGA	5700
5641	Db	ATTTTCATCAGTGGGATACATATTTTGGCGGGCTGTCAACGCTGCTGTGTAAACCCCGCCA	5700
5701	Qy	TAGCATCATTGATGGCATTTTACAGTCTTCTATCACTAGCCCGCTCACACCCAAAAACCCC	5760
5701	Db	TTGCTTCATTGATGGCTTTTACAGTCCCGTCAACAGCCCACTAACCACTGGGCCAAACCC	5760
5761	Qy	TCCTGTTTAACTCTTTGGGGGATGGGTGCTGCCAACTCGCTCTCCAGCGCTGCGT	5820
5761	Db	TCCTTTTCAACATATTGGGGGGTGGGTGGCTGCCAGCTCGCGCGCCCCGCTGCGGCTA	5820
5821	Qy	CAGCTTTCTGTGGGCGCGGCATCGCCGAGCGGCTGTTGGCAGCATAGGCTTGGGGAAG	5880
5821	Db	CCGCTTTTGTGGCGCTGGCTTATAGCTGGCGCCGCCATTCGGCAGCTTGGACTGGGGAAG	5880
5881	Qy	TGCTCGTGGACATCTTTCGGGGCTATGGGGCAGGGGTAGCCGCGCACTCGTGGCCCTTTA	5940
5881	Db	TCCTCTGTGGACATCTTTCAGAGGTATGGCGGGCGTGGGGAGCTCTGTGTAGCATTC	5940
5941	Qy	AGGTCATGAGCGGCGAGGTCCCTCCACCGAGACCTGTCAACTTACTCCCTGCCATCC	6000
5941	Db	AGATCATGAGCGTGAAGTCCCCCTCCACGAGGACCTGTGTCATCTGCTACCCGCCATCC	6000
6001	Qy	TCTCTCTGTGGTCCCTGGTCTGCGGGTCTGTGGCAGCAATACTGCGTTCGCGCAGCTGG	6060
6001	Db	TCTCGCTTGGAGCCCTTGTAGTCCGTTGTGTCGTGCGCAGCAATCTGCGCGCGCACGTTG	6060
6061	Qy	GCCCGGAGAGGGGCTGTCAGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGGGGTA	6120
6061	Db	GCCCGGCGAGGGGCGAGTGCAATGGATGAACCGGCTAATAGCTTCGCCCTCCCGGGGA	6120
6121	Qy	ACCACTGCTCCCTTCACCATATGTGCTCGAGAGCGCTGCAGCACGTGTCTCACTCAGA	6180
6121	Db	ACCATGTTTCCCCACGCACTAGTGCCGGAGAGCGATGCAGCGCCGCGTCACTGCCA	6180
6181	Qy	TCCTCTTAGCCCTTACCATCATCAACTGCTGAAGCGGCTCCACCACTGATTAATGAGG	6240
6181	Db	TACTCAGCAGCCTCACTGTAAACCAAGCTCTCTAGGCGCACTACATCAGTGGATTAAGCTCG	6240
6241	Qy	ACTGCTTAGCCCATGCTCCGGCTCGTGGCTTAAGGATGTTTGGGATTTGGATATGCACGG	6300
6241	Db	AGTGTAACACTTCCATGCTCCGGCTCTTGGCTTAAGGACATCTGGGACTGGATATCGCAGG	6300
6301	Qy	TGTTGACTGACTTCAAGACCTGGCTCCAGTCCAAACTCTCTGCCGCGTTTACCGGAGTCC	6360
6301	Db	TGCTGAGCGACTTTTAAGACCTGGCTGAAGCCAAAGCTCATGCCACAACCTGCTGGGATTC	6360
6361	Qy	CTTTCCTGTCAATGCCAAACGCGGGTACAAAGGAGTCTCGCGGGGGAGCGGCATCATGCAA	6420
6361	Db	CCTTGTCTCTGCAGCGCGGGTATATAGGGGGTCTGCGCAGGAGAGCGGCATTAATGACA	6420
6421	Qy	CCACTGCGCATGCGGAGCACAGATCCGCGGACATGTCAAAAACGTTTCCATAGGATCG	6480
6421	Db	CTCGTGGCACTGTGGAGCTCAGATCACTGACATGTCAAAAACGCGACGATGAGGATCG	6480

Qy 6481 TAGGSCCTAGAACCTGACGCAACACGTTGCGACGGAAACGTTCCCATCAACGCATACACCA 6540
Db |||||
Qy 6481 TCGGTCCTTAGNCCCTGCGAGAACATGTGGAGTGGACGTTCCCATTAACGCTACACCA 6540
Db |||||
Qy 6541 CGGACCTTTGACACCTCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGGTGGCTG 6600
Db |||||
Qy 6541 CGGCGCCTGTACTCCCTTCTCGCGCGAACTATAAGTTGCGCGTGTGGAGGGTGTCTG 6600
Db |||||
Qy 6601 CTGAGAGTACGTGAGAGTTAGCGGTGGGGATTTCCACTAGTGACGGCATGACCA 6660
Db |||||
Qy 6601 CAGAGGAATACGTGGAGATAAGCGGGTGGGGACTTCCACTAGTATCGGGTATGACTA 6660
Db |||||
Qy 6661 CTGACAAAGTAAAGTGCCTACGCTCCGAGGTTCCGGCCCCCGAAATCTTACGCGAGGTGGATG 6720
Db |||||
Qy 6661 CTGACAAATCTTAATGCGCGTGCAGATCCCATGCCCGGAATTTTACAGAAATGGAGC 6720
Db |||||
Qy 6721 GAGTCCGTTGACAGGTACGCTCCGGCGTGCAGAACTCTTCTACGGGAGGACGTCAGT 6780
Db |||||
Qy 6721 GGGTCCGCTACATAGGTTTGGCGCCCTTGCAGCCCTTGTGCGGGAGGAGGTATCAT 6780
Db |||||
Qy 6781 TCCAGGTCGGGCTCAACCAATACTTGGTGGGTGCGAGCTCCCATGCGAGCCGACCGG 6840
Db |||||
Qy 6781 TCAGAGTAGGACTCACAGATACCGGGTGGGTGCGCAATTAACCTTGCAGAGCCGAAACCGG 6840
Db |||||
Qy 6841 AGCTAACAGTCTTACTTCCATGCTCACGATCCCTCCCAATTAACAGCAGAGACGGCTA 6900
Db |||||
Qy 6841 AGCTAGCGGTTGACGTCCTACGCTCATGATCCCTCCCATATAACAGCAGAGGGCGCG 6900
Db |||||
Qy 6901 AGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTTAGCCAGCTCAFCAGCTAGCCAGTTGT 6960
Db |||||
Qy 6901 GGAGAAGGTTGCGGAGAGGGTCACCCCTTCTATGGCCAGCTCCTCGGCCAGCCAGCTGT 6960
Db |||||
Qy 6961 CTGCGCCTTTTGAAGGGGACATGCACTA CCCACCATGATCTCCCGGAGCGTGACCTCA 7020
Db |||||
Qy 6961 CCGCTCCATCTCTCAAGGCAATTGTCACCGCCCAACCATGACTCCCTGACCGCGAGCTCA 7020
Db |||||
Qy 7021 TCGAGGCCAACCTTGTGGCGGAGGAGATGGCGGAAACATCACTCGCTGGAGTGCAG 7080
Db |||||
Qy 7021 TAGAGGCTAACTCTCTGTGGAGGAGGAGATGGCGGCAACATCAACAGGTTGAGTCA 7080
Db |||||
Qy 7081 AGAATAAGGTAGTAATCTTGACTCTTTTCGAAACCGCTTCA CGCGGAGGGGATGAGAGG 7140
Db |||||
Qy 7081 AGAACAAAGTGTGATCTGGACTCTTCGATCCGTTGTGGCAGAGGAGATGAGCGG 7140
Db |||||
Qy 7141 AGATATCCGTCCGGGGAGATCTTGCGAAATCAGAGAAAGTTCCCTCAGGGTTGCCCA 7200
Db |||||
Qy 7141 AGGTCTCCGTACCCGAGAAAATTCTGCGAAGTCTCGGAGATTCCCGCGGCCCTGCCCG 7200
Db |||||
Qy 7201 TATGGGCA CGCCGGACTACAATCTTCCACTGCTAGAGTCTGGAAGGACCGGACTAGC 7260
Db |||||
Qy 7201 TTTGGCGCGCGGACTACAACCCCTAGTGAAGACGTGGAAAGCTGACTAGC 7260
Db |||||
Qy 7261 TCCCTCCGGTGTACACGATCCCATTTGCCACCTACCAAGGCTCTCCAAATACACCTC 7320
Db |||||
Qy 7261 AACCACTCTGTGTCATGCTGCGCGCTACCACTCCAGGTCCTCTCTGTGCCTCCG 7320
Db |||||
Qy 7321 CACGGAGAAAGGAGCGGTTGTCTGACAGAAATCCAATCCAATGTGTCTTCTGCTTGGCGGAGC 7380
Db |||||
Qy 7321 CTCGGAAGAGGAGTACGGTGTCTTCCACGAAATCAACCCCTACTGCTTGGCGGAGC 7380
Db |||||
Qy 7381 TCGCCACTAAGACCTTCGGTAGCTCCGATCTGCGCGGTGATAGCGGACGGGACCG 7440
Db |||||
Qy 7381 TTGGCCACCAAAAGTTTGGCAGCTCTCAACTTCCGGCAATACGGGACAAATATGACAA 7440
Db |||||
Qy 7441 CCGCTTCTGACCTGGCCTCCGACGAGCGGTGACAAAGGATCCGACGTTGAGTGTACTCTCT 7500
Db |||||
Qy 7441 CATCTCTGAGCCCGCCCTTCTGGCTGCCCGCCGACTCCGACGTTGAGTCTTATTCTT 7500
Db |||||
Qy 7501 CCAATGCCCCCTTGAAGGGAGCGCGGGGACCCCGATCTCAGCAGCGGCTTCTGGTCTA 7560
Db |||||
Qy 7501 CCAATGCCCCCTTGAAGGGAGCGCGGGATCCCGATTTTACGACCGGGTATGCTGCA 7560
Db |||||

Qy 7561 CCGTGAAGTGAAGGGCT---AGTGAAGATGTCTGTCTCTCAATGTCTTATACGTGA 7617
Db |||||
Qy 7561 CCGTCACTAGTGTGGGCGACACGAAAGATGTCTGTCTCTCAATGTCTTATACCTGA 7620
Db |||||
Qy 7618 CAGCGCCCTGATCA CGCCATGCGCTGCGGAGGAAAGTAAGCTGCCCATCAACCCGTTGA 7677
Db |||||
Qy 7621 CAGCGCACTCGTCA CCGCGTGCCTGCGGAAGAACAAAACTGCCCATCAACGCACTGA 7680
Db |||||
Qy 7678 GCAACTCTTGTCTGCTGACCAACATGCTCTACGCGCAACATCCGCGAGCGCAAGCC 7737
Db |||||
Qy 7681 GCAACTCTTGTCTAGCCCATCA CAATCTGTATATTCACCACTTCAGCAGTGTGCTTGC 7740
Db |||||
Qy 7738 TCGCGCAAGAAAGGTCA CACTTTTACAGATTTGCAAGTCTCTGATGATCATTTACCGGAGC 7797
Db |||||
Qy 7741 AAAGCGAAGAAAGTCA CACTTTTGACAGACTGCAAGTTCTGGACAGCAATTTACCAGGAGC 7800
Db |||||
Qy 7798 TACTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTTAAGGCTAAGCTTCTATCTATAGAG 7857
Db |||||
Qy 7801 TGCTCAAGGAGGTCAAAGCAGCGCGTCAAAGTGAAGGCTAACTTGTCTATCCGTAGAG 7860
Db |||||
Qy 7858 AGGCTCAAGCTGACGCCCCC CACTTCGCGCAAAATCCAAATTTGGCTATGGGCAAGG 7917
Db |||||
Qy 7861 AAGCTTCAGCTGACGCCCC CCACTTCAGCAGAAATCCAAAGTTTGGCTATGGGCAAGG 7920
Db |||||
Qy 7918 ACCTCCGGAACCTATCCAGCAGGCGCTTAA CCACTCCGCTCCGTGTGGAGGACTTGC 7977
Db |||||
Qy 7921 AGTTCGCTGCCATGCGCAGAAAGCGCTAG CCCCACATCAACTCTCGTGTGGAAGACCTTC 7980
Db |||||
Qy 7978 TGAAGACACTGAAACACCAATTGACACCACTCATGGCAAAAGTGAGGTTTTCTGCG 8037
Db |||||
Qy 7981 TGAAGACAGTGTAA CCAATAGACACTATCATCATGGCCAAAGACGAGGCTTCTTCTGCG 8040
Db |||||
Qy 8038 TCCAAACAGAGAGGAGCGCGCAAGCCAGTTCGCTTATCGTATTCAGACCTGGAG 8097
Db |||||
Qy 8041 TTAGCTGTGAAGGGGGTCTGAAGCCAGTCTGCTCATCGTGTCCCCGACCTGGCG 8100
Db |||||
Qy 8098 TTAGCTGTATGCGAAGATGGCCCTTTACGACGCTGCTCCACCCCTTCTCAGGCGCTGA 8157
Db |||||
Qy 8101 TCGCGTGTGCGAAGATGGCCCTGTACGAGCTGTTAGCAAACTCCCCCTGGCGCTGA 8160
Db |||||
Qy 8158 TGGGCTCTCATACGGAATTCATCTCC CCAAGCAGCGGCTCGAGTTCTCTGTGAATA 8217
Db |||||
Qy 8161 TGGGAAGCTCTACGGAATTCATCTCA CCGAGACAGCGGTTGAATTCCTCGTGCAG 8220
Db |||||
Qy 8218 CCGGAATCAAGAAATGCGCTTATGGCTTCTCATATGACCCGCTGTTTTCAGCTCA 8277
Db |||||
Qy 8221 CGTGAAGTCCAAAGACCCCGATGGGTTT CCGGTATGATACCCGCTGTTTGACTCA 8280
Db |||||
Qy 8278 CGGTCACTGAGAGTGACATTCGTTGAGGAGTCAA TTTACCAATGTTGTGACTTGGCCC 8337
Db |||||
Qy 8281 CAGTCACTGAGAGCGACATCCGTA CCGGAGGAGGCAATTTACCAATGTTGTGACTGGACC 8340
Db |||||
Qy 8338 CCGAGGCGCAGACGCGCATTAAGGTGCGTCA CAGAGCGGCTTTACATCGGGGGTCCCCCTGA 8397
Db |||||
Qy 8341 CCCAAGCCCGGTGCCCATCAAGTCCCTCACTAGAGGCTTTATGTTGGGGGCCCTCTTA 8400
Db |||||
Qy 8398 CTAACTCAAGAGGACAGAACTGCGGTTATCG CCGGTGCGCGCAAGTGGCGTGTGACGA 8457
Db |||||
Qy 8401 CCAATTAAGGGGGGAAACTGCGGCTATCG CAGGTGCGCGGAGCGGCTACTGACAA 8460
Db |||||
Qy 8458 CTAGTCCGGTAATACCTCTACATGTTACTTTGAAGGCGCACTGCGAGCTGTGCGAGTGC 8517
Db |||||
Qy 8461 CTAGTGTGGTAA CACCTCACTTGTACTCAAGGCCCGGCGAGCCGCTCGAGCCGAG 8520
Db |||||
Qy 8518 AGTCCAGGACTGACGATGCTCGTGAACCGGAG CAGCTTGTGTTTCTGTGTAAGCGG 8577
Db |||||
Qy 8521 GGCTCCAGGACTGCA CCAATGCTCGTGTGCGCAGCACTTAGTGTGTTATCTGTGAAAGTG 8580
Db |||||
Qy 8578 CCGGAA CCGCAGGAGGTCGCGGCCCTTACGAGCTTCACGAGGCTATGACTAGTAT 8637
Db |||||
Qy 8581 CCGGGGTCAGGAGGACCGCGCGAGCTTGAGAG CTTTACGAGGCTATGACAGGTA 8640
Db |||||
Qy 8638 CCGCCCCCGGGGATCCGCCCAACCAAGAAATACGACCTGGAGCTGATTAACATCATGTT 8697
Db |||||

Db 8641 CGCCCCCCCCGGGACCCCCCAACAGAAATACGACTTGGAGCTTATAACATCATGCT 8700
Qy 8698 CTTCCAAATGTGTCAGTCGCGCAGATGTCATCTGGCAAAAGGGTATATCTACTCCACCGTG 8757
Db 8701 CTTCCAAAGTGTGTCAGTCGCGCAGCAGCGGCTGGAAAAGGGTCTACTACTTACCGTG 8760
Qy 8758 ACCCCACACCCCTTTCAGCGGCTGCGGGGAGACAGCTAGACACACTCCAAATCACT 8817
Db 8761 ACCCTACACCCCTTTCAGCGGCTGCGGGGAGACAGCTAGACACACTCCAAATCACT 8820
Qy 8818 CTTGGCTAGCAATATCATGTATGCGGCCACCTATGCGCAAGATGATCTGATGA 8877
Db 8821 CTTGGCTAGCAATATCATGTATGCGGCCACCTATGCGCAAGATGATCTGATGA 8880
Qy 8878 CTCACATTTTCTCATCTTCTAGCTCAAGAGCAACTTGAAGAGCCCTGGATGTGAGA 8937
Db 8881 CCATTTCTTAGGCTCTCATAGCAGGATCAGCTTGAACAGGCTCTTAACCTGTGAGA 8940
Qy 8938 TCTACGGGGCTTGTACTCTCAATGAGCCACTTGAACCTTACCTCAGATCAATGAAAGCTCC 8997
Db 8941 TCTACGGGGCTTGTACTCTCAATGAGCCACTTGAACCTTACCTCAGATCAATGAAAGCTCC 9000
Qy 8998 ATGGTCTTAGGCTTACTCTCAATGAGCCACTTGAACCTTACCTCAGATCAATGAAAGCTCC 9057
Db 9001 ATGGCTCAGCGCAATTTTCACTCCACAGTTACTCTCAGGTTAGTCAATAGGGTGGCG 9060
Qy 9058 CATGCTCAGAAACTTGGGGTACACCTTGGCAACTCGAGACATCGGGCCAGAAAGTG 9117
Db 9061 CATGCTCAGAAACTTGGGGTACACCTTGGCAACTCGAGACATCGGGCCAGAAAGTG 9120
Qy 9118 TCCGCGCTAAGCTACTCTCCAGGGGGAGGGCCGCCACTTGTGGCAGATACCTTTTA 9177
Db 9121 TCCGCGCTAAGCTACTCTCCAGGGGGAGGGCCGCCACTTGTGGCAGATACCTTTTA 9180
Qy 9178 ACTGGGAGTAAAGCAAGCTTAACTCACTCAATCCCGCCGCTCCAGCTGACT 9237
Db 9181 ACTGGGAGTAAAGCAAGCTTAACTCACTCAATCCCGCCGCTCCAGCTGACT 9240
Qy 9238 TGTCTGCTGCTTCTGCTGCTTACAGCGGGGAGACATATATCAAGCTGCTCTGCTG 9297
Db 9241 TGTCCGGTGTGCTTACCGCTGCTTACAGCGGGGAGACATATATCAAGCTGCTCTGCTG 9300
Qy 9298 CCGGACCCGCTGCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9357
Db 9301 CCGGACCCGCTGCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9360
Qy 9358 TGTCTCCCAACCGATGAACGGGGAGCTAAACACTCCAGGCTTAAAGCCATTTCTGTTTT 9417
Db 9361 TCTCTCCCAACCGGTAAGGTTGGGGTAAACACTCCGCGCTCTTAAAGCCATTTCTGTTTT 9420
Qy 9418 -----TTTCT 9455
Db 9421 TTT 9480
Qy 9456 TTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 9513
Db 9481 TCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 9540
Qy 9514 CCCTAGTCAACGGCTAGCTGTGAAGGTCCTGAGCCGATGACTGCGAGAGTGTCTGATA 9573
Db 9541 CCCTAGTCAACGGCTAGCTGTGAAGGTCCTGAGCCGATGACTGCGAGAGTGTCTGATA 9600
Qy 9574 CTGGCTCTCTGCAGATCATGT 9595
Db 9601 CTGGCTCTCTGCAGATCATGT 9622

RESULT 14
US-09-238-076-1
; Sequence 1, Application US/09238076
; Patent No. US20020102540A1
; GENERAL INFORMATION:

APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-238-076-1

Query Match 65.2%; Score 6254.8; DB 3; Length 9646;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 7558; Conservative 0; Mismatches 2037; Indels 51; Gaps 3;
Qy 1 GCCAGCCCCCTGATGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAATCTG 60
Db 1 GCCAGCCCCCTGATGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAATCTG 60
Qy 61 TCTTACGCAAGAAAGCGCTAGCCATGGCGTTAGTATGAGTGTGTCAGCTCCAGGAC 120
Db 61 TCTTACGCAAGAAAGCGCTAGCCATGGCGTTAGTATGAGTGTGTCAGCTCCAGGAC 120
Qy 121 CCCCCCTCCGGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTCGAC 180
Db 121 CCCCCCTCCGGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTCGAC 180
Qy 181 GACGACCGGGTCTCTTTCTTGGATCAACCCCGCTCAATGCTGGAGATTGGCGTCCCCC 240
Db 181 GACGACCGGGTCTCTTTCTTGGATCAACCCCGCTCAATGCTGGAGATTGGCGTCCCCC 240
Qy 241 GCGAGACTGTAGCCGAGTGTGTGGGTGCGAAAGCCCTTGTGGTACTGCTGTATAGG 300
Db 241 GCAAGACTGTAGCCGAGTGTGTGGGTGCGAAAGCCCTTGTGGTACTGCTGTATAGG 300
Qy 301 GTGCTTGGAGTGTGCGGGAGGTTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360
Db 301 GTGCTTGGAGTGTGCGGGAGGTTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360
Qy 361 CTCAAGAAAAACCAAAACGTAACCAACCGTCCGCCACAGGACGTCAGTTCCCGGGCG 420
Db 361 CTCAAGAAAAACCAAAACGTAACCAACCGTCCGCCACAGGACGTCAGTTCCCGGGCG 420

QY 421 GTGGTCAGATCGTTGGTGGAGTTTACCTTTGTCGCGAGGGGCCCCAGGTGGGTGTC 480
Db 421 GGGTCAGATCGTTGGTGGAGTTTACTTTGTTGCGCGCAGGGGCCCTAGATTGGGTGTC 480
QY 481 GCGCACTAGGAAGCTTCCGAGCGGTGCAACCTCTGTGGGAAGGCGCAACCTATCCCAA 540
Db 481 GCGCACTAGGAAGCTTCCGAGCGGTGCAACCTCTGTGGGAAGGCGCAACCTATCCCAA 540
QY 541 AGGCTCGCGGACCCGAGGCGAGGGCTGGGCTCAGCCCGGTACCCCTTGGGCCCTCTATG 600
Db 541 AGGCACTGCGGCCCGAGGCGAGGACCTGGGCTCAGCCCGGTACCCCTTGGGCCCTCTATG 600
QY 601 GCAATCAGGGCTGCGGGTGGGAGGATGGCTCTCTCAACCGCGGCTCCGCGCTAGTT 660
Db 601 GCAATCAGGGTGGCGGTGGGAGTGGCTCTCTCTCCCGTGGCTCTGGGCTAGCT 660
QY 661 GGGGCCCAAGGACCCCGGCGTAGGTTCGCGTAATCTGGGTAAAGTCAATCGATACCCCTTA 720
Db 661 GGGGCCCAAGGACCCCGGCGTAGGTTCGCGTAATCTGGGTAAAGTCAATCGATACCCCTTA 720
QY 721 CATGGGCTTCGCGATCTCATGGGGTACATTCGCTCGTCGGCGGCCCTTAGGGGCG 780
Db 721 CGTGGGCTTCGCGGACCTCATGGGGTACATACCGCTCGTCGGCGGCCCTCTTAGGGGCG 780
QY 781 CTGCCAGGCTTGGCACACGGGTCCGGGTTCTGGAGGACGGCGTGAATATGCAACAG 840
Db 781 CTGCCAGGCTTGGCGCATGGCTCGGGTCTCGAGAGCGGCGTGAATATGCAACAG 840
QY 841 GGAATTCGCCGGTTCCTTTCTCTATCTTCTCTCTTGGCTCTGCTGCTGCTTTGACCA 900
Db 841 GGAACCTTCGTTGCTCTTTCTCTATCTTCTCTCTTGGCCCTGCTCTTGGCTGACTG 900
QY 901 TCCGAGCTTCGCTTATGAAGTGGCGACAGTGTCCGGATATACCATGTCAAGACGACT 960
Db 901 TCGCGCTTTCGAGCTACCAAGTGGCGAATTCCTCGGGGCTTTACCATGTCAAGACGACT 960
QY 961 GCTCAACTCAAGCATTTGTATGAGGACGAGCGATGATCATGATATCTCCGGGTGCG 1020
Db 961 GGCCTAACTCGAGTATTTGTACGAGGCGGCGATGCCATCTGCACACTCGGGGTGTG 1020
QY 1021 TGCCTGTGTTTCAAGAGGTAAACGCTCCGTTGCTGGGTAGCGTCACTCCCAAGCTCG 1080
Db 1021 TCCCTTGCTTCGAGGGTAAACGCTCGAGGTGTGGGTGGCGGTGACCCCAAGCTGG 1080
QY 1081 CCGCCAGGAATGCCAGCTGCCCATACGACAAATAGCAGCCACGCTCGACTTGTCTGTTG 1140
Db 1081 CCACCAAGGACGGCAAACTCCCCAAGCAGCTTCGAGCTCATATCGATCTGCTTGTGCG 1140
QY 1141 GGACGCTGCTTTCTGCTCCGCTATGTACGTGGGGATCTCTCGCGATCTATTTTCTCTCG 1200
Db 1141 GGAGGCCAACCTCTGCTCGGCCCTACGTGGGGACCTGTGCGGGTCTGCTTTCTTGTG 1200
QY 1201 TCTCCAGCTGTCACTTCTCGCTCGCCGATGAGACAGTGCAGGACTGCAACTGCT 1260
Db 1201 TTGGTCAACTGTTTACTTTCTCTCCAGCGCCACTGGACGACGCAAGACTGCAATTTGT 1260
QY 1261 CAATCTATCCGGCCATGTATCAGGTCAACCGATGCGCTTGGGATATGATGAAGTGGT 1320
Db 1261 CTATCTATCCGGCCATATAACGGGTCAATCGATGGCATGGGATATGATGAAGTGGT 1320
QY 1321 CACCTACAAAGCCCTAGTGGTGTGCGAGTGTCTCGGATCCCAAGCTGCTGGGACA 1380
Db 1321 CCCCTACGGAGGTTGGTGGTATGCTCAGCTGCTCGGATCCCAAGCCATCATGGACA 1380
QY 1381 TGGTGGCGGGGCCCACTGGGAGTCTCGCGGCTTCCCTACTATTTCCATGGTAGGGA 1440
Db 1381 TGATCGCTGGTGTCTACTGGGAGTCTCGCGGCAATAGCGTATTTCTCATAGTGGGGA 1440
QY 1441 ACTGGGCTAAGGTTCTGATTTGGCGCTACTCTTTGCCGCGGCTTGACGGGAGACCCACA 1500
Db 1441 ACTGGGCGAAGGTCCTGGTAGTGTCTGCTATTTGCGCGGCTGCAAGCGGAAACCCACG 1500
QY 1501 CGACGGGGAGGGTGGCGGCCACACCACTCGGGGTTACGTCCTCTTTTCTCATCTGGGG 1560

Db 1501 TCACGGGGGAAGTGC CGCGCCGACCAACGCTGGGCTTGTGGTCTCTTACACAGCGC 1560
QY 1561 CGTCTCAGAAAATCCAGCTTGTGAATACCAACGCGAGCTGCGCATCAACAGGAGTGC 1620
Db 1561 CCAAGCAGAACATCCAACTGATCAACACCAACGCGAGTTGGCATCAATAGACGGGCT 1620
QY 1621 TAAATTCGAATGACTCCCTCCAAACTGGGTTCTTTGCGCGCTGTTTATCGCACACAAGT 1680
Db 1621 TGAATGCAATGAAGCTTAAACACCGCTGGTTAGCAGGCTCTTCTATCAGACAAAT 1680
QY 1681 TCAACTCGTCCGGTGC CGGAGCGCATGCGAGCTGCGGCCCACTGACTGTGTGCGCC 1740
Db 1681 TCAACTCTTCAGGCTGTCTGAGAGGTGGCAGCTGCGACGCTGCCAGCTTACCGATTGCGCC 1740
QY 1741 AGGGTGGGGCCCATCACCTATATAAGCTTAACAGCTCGGATCGGATCAGAGGCTTATGCT 1800
Db 1741 AGGCTGGGGTCTTATCAGTTATGCCAAACGGAAGCGGCTCGAGAAACGCGCCCTACTGCT 1800
QY 1801 GGCAATACGCGCTCGACCGTGTGCTGTACCCGCGTGCAGGTGTGTCAGGTGTGTCAGTGT 1860
Db 1801 GGCACTACCTCCAAAGACCTTGTGGCATTTGCCCGCAAGAGAGCTGTGTGCGCCGGTAT 1860
QY 1861 ATTGTTTACCCCAAGCCCTGTTGTGTGGGACCAACCGATTCGTTCCGGTGTCCCTACGT 1920
Db 1861 ATTGCTTCACTCCAGCCCGTGTGTGGGAACGACGACAGGTCCGGCGCGCTACCT 1920
QY 1921 ATAGTGGGGGAGAAATGACACAGCGTGTATCTCTCAACACACGCGTCCGCGACAAAG 1980
Db 1921 ACAGCTGGGGTGAATGATACGGATGTCTTCGTCTTAAACACACGAGGACACGCTGG 1980
QY 1981 GCAACTGGTTCGGCTGTACATGGATGAATAGTACTCGGTTTCACTAAGACGTGCGGAGTC 2040
Db 1981 GCAATTTGGTTCGGTGTACTTGGATGAATCACTGGATTCACCAAGTGTGCGGAGCGC 2040
QY 2041 CCCGTTGTAAATCGGGGGGTGGTAAACCGACCTTTGATCTGCCCAACGAGTGTCTTCC 2100
Db 2041 CCCCTTGTGTATCGGAGGGTGGGCAACACACTTGTCTGCTGCCCACTGATTGTTTCC 2100
QY 2101 GGAAGCACCCGAGGCTACTTTACAAAATGTGCTCGGGCCCTGGTTGACACCTAGGT 2160
Db 2101 GCAAGCATTCGGAAGCACATCTCTCGGTGCGGCTCGGTCTCGGATTAACCCAGGT 2160
QY 2161 GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACCTCAATTTTCCATCT 2220
Db 2161 GCATGTGCACTACCCGTATAGGCTTTGGCACTATCTTGTACCATCAATTTACCATAT 2220
QY 2221 TTAAGTTAGGATGTATGTGGGGCGTGGAGCACAGGCTCAATTCGCGCATGCAATTGGA 2280
Db 2221 TCAAAGTCAGGATGTACGTGGAGGGGTGAGACACAGGCTGGAAGCGGCTGCAACTGGA 2280
QY 2281 CTCGAGGAGCGCTGTAACTTGGAGACAGGATAGGTAGAACTCAGCCGCTGCTGTC 2340
Db 2281 CGCGGGCGAAACGCTGTGATCTGGAAGACAGGAGCAGGTCCGAGCTCAGCCCATTTGCTG 2340
QY 2341 TGTCTCAACACAGGTGGCAGATCTGCCCTGTCTTTCACCACTACCGCTTATCCA 2400
Db 2341 TGTCCACACACAGTGGCAGGCTCTTCCGTGTTCTTTCACGACCTTGCAGCCTTGCTCA 2400
QY 2401 CTGGTTTGTATCCATCTCCATCAGAAACATCTGTGGAGCTGSCAAATACCTGTACGGT 2460
Db 2401 CCGGCTCATCCACTCCACAGAAACATTTGTGGAGCTGACGTACTTTGTACGGGTAGGT 2460
QY 2461 CAGCGTTTGTCTCTTTGCAATCAATGGGAGTACATCTCTGTTGCTTTTCTCTCTCTG 2520
Db 2461 CAAGCATCGCTCTCGGCCCAATTAAGTGGGAGTACGTGTTCTCTGTTCTCTCTCTG 2520
QY 2521 CAGAGCGCGGTGTGTCCTGTTGTGGATGATGCTGCTATAGCCAGCCAGGCTGAGGCG 2580
Db 2521 CAGAGCGCGGTGTGCTCTGTTGTGGATGATGTTACTCATATCCACAGGAGGCGG 2580
QY 2581 CTTTAGAGAACTTGGTGGTCTCAATGCGGCTCCGTGGCGGAGCGCATGTTCTTCT 2640

QY 4801 GCGAGGTAGAACTGCGAGGGTAGAGTGGGCATCTACAGGTTTGTGACTCCAGAGAAAC 4860
DB 4801 GCCGGGCGAGGATGGCAGAGGGGAAGCAGGCACTACAGATTGTGGCACCGGGGAGC 4860
QY 4861 GGCCTTCGGGCATGTTGCAATCTTCGGTCTGTGTGAGTGTATGACCGGGGCTGTGCTT 4920
DB 4861 GCCCTTCGGGCATGTTGCAACTCGTCCGTCTGTGTGAGTGTATGACCGGGGCTGTGCTT 4920
QY 4921 GGTATGACTCAGCCCGCTGAGACTCGGTAGGTTCGGGGCTTACCTAATACACAG 4980
DB 4921 GGTATGACTCAGCCCGCTGAGACTCAGGTAGGTTCGGGGCTTACCTAATACACAG 4980
QY 4981 GGTTCGGCTTCGCCAGGACCATCTGGAGTTCTGGGAGAGGCTTTCACAGGCTTCACCC 5040
DB 4981 GGTTCGGCTTCGCCAGGACCATCTTGAATTTTGGAGGGGCTCTTTACGGGCTCACTC 5040
QY 5041 ACATAGATGCCCATCTTCGTGCCAGACTAAACAGGCAGGAGACAACTTCTTACCTGG 5100
DB 5041 ATATAGATGCCCATCTTATCCAGACAAAGCAGAGTGGGAGAACTTCTTACCTGG 5100
QY 5101 TGGCATATCAAGCTACAGTGGCGCGAGGGCTCAAGCTCCACCTCGTGGGACAA 5160
DB 5101 TAGCGTACCAAGCCACCGTGTGCTAGGGCTCAAGCCCTCCGCCATCGTGGGACCA 5160
QY 5161 TGTGGAAGTGTCTCATACGGCTGAAACCTTACACTGACGGGCAACACCCCTGCTGTATA 5220
DB 5161 TGTGGAAGTGTGTGATTCGCCCTTAAACCCACCTTCATGGGCCAACACCCCTGCTATA 5220
QY 5221 GGTAGGAGCGCTCAAAATAGAGTCATCTCCACACACCCCAATACTAAATACATATGG 5280
DB 5221 GACTGGGCGCTGTTCAGAATGAAGTCACCTGACGACCCCAATCACCAATACATATGA 5280
QY 5281 CATGATGCTGCGCTGACCTGGAGGTGCTCATAGCACCTGGGTGCTGTAGCGGAGTCC 5340
DB 5281 CATGATGCTGCGCGACCTGGAGGTGCTCATAGCACCTGGGTGCTGTAGCGGAGTCC 5340
QY 5341 TTGACAGTTTGGCGCATCTCCCTGACGACAGGAGTGTGTGATGTTGGGACAGATCA 5400
DB 5341 TGGCTGCTTGCGCGGCTATTGCCCTGTCAACAGGCTGCTGTGTCATAGTGGGACGATG 5400
QY 5401 TCTTGTCCGGGAAGCCAGCTGCTGTTCCCGACAGGGAAGTCTCTACAGGAGTTTCGATG 5460
DB 5401 TCTTGTCCGGGAAGCCGCAATTATACCTGACAGGAGGTTCTCTACAGGAGTTTCGATG 5460
QY 5461 AGATGAGAGAGTGTGCTCACAACCTTACATCGAGCAGGGAATGACGTCGCCGAGC 5520
DB 5461 AGATGAGAGAGTGTCTCAGCACTTACCGTACATCGAGCAAGGGAATGATGCTCGCTGAGC 5520
QY 5521 AATTCAAGCAAAAGGCGCTCGGTTGTTGCAAAAGCCGACCAAGCAAGCGAGGCTGCTG 5580
DB 5521 AGTTCAAGCAGAAAGGCGCTCGGCTCTGCGCTCTGACACCGGCTCCGCCAGGAGTTATCA 5580
QY 5581 CTCCGTTGGTGGAGTCAAGTGGGAGCCCTTGAGACCTTTGGGCGAAGCACAATGTGGA 5640
DB 5581 CCCCTGCTGTCCAGACCAACTGGCAGAACTCGAGGCTCTTCTGGGCGAAGCACAATGTGGA 5640
QY 5641 ATTTCATAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGTGAACCCCGGA 5700
DB 5641 ATTTCATAGTGGGATACAAATCTTGGCGGGGCTGTCAACGCTGCTGTGTAACCCCGCA 5700
QY 5701 TAGCATCATGATGCAATTTACAGCTTCTATCACTAGCCCGCTCACCAACCCCAACACCC 5760
DB 5701 TTGCTTCATGATGGCTTTTACAGCTGCGGTCAACAGCCCACTTAACCACTGCGCCAAACCC 5760
QY 5761 TCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCCCAATCGCTCTCTCCAGCGCTGGT 5820
DB 5761 TCTCTTCAACATATTGGGGGGGTGGGTGGCTGCCCCAGCTCGCGCCCGCGTGCCTA 5820
QY 5821 CAGCTTTCGTGGCGCGCGCATCGCCGAGCGGCTGTTGGAGCATAGGCCCTTGGGAAG 5880
DB 5821 CCGCTCTTGTGGCGCTGCTTAGCTGGCGCGGCCATCGGCAGCGTTGACCTGGGGAAG 5880
QY 5881 TGCTCGTGACATCTTGGCGGGCTATGGGCGAGGGGTAGCGCGCACTCGTGGCTTTA 5940

DB 5881 TCCTCGTGGACATTTCTTGAGGGTATGGCGCGGCGTGGCGGAGCTCTTGTAGCAATCA 5940
QY 5941 AGTTCATGAGCGGCGAGGTGCCCTCCACAGAGCACTGGTCAACTTACTCCCTGCCATCC 6000
DB 5941 AGATCATGAGCGGTGAGGTCCCTCCACAGAGCACTGGTCAACTCTGCTGCCCGCATCC 6000
QY 6001 TCTCTCTGTCCTGCTGCTGCGGGTCTGTGCGCAGCAATACTCGCTCGGCACTGG 6060
DB 6001 TCTCGCTTGAGCCCTTGTAGTGGTGTGTCTGCGCAGCAATACTGCGCGGCACTGG 6060
QY 6061 GCCCGGAGAGGGGCTGTGCACTGAGTGAACCGGCTGATAGCGTTTCGTTTCGGGGGTA 6120
DB 6061 GCCCGGAGAGGGGCTGCAATGGATGAACCGGCTAATAGCCTTCCCTCCCGGGGGA 6120
QY 6121 ACCAGTCTCCCTTACGCACTATGTGCTGAGAGCGAGCTGCGAGCACTGCTCAG 6180
DB 6121 ACCATGTTTCCCGACCACTAGCTGCGGAGAGCGATGAGCGCCCGCTCACTGCCA 6180
QY 6181 TCCTCTCTAGCCTTACCCTCACTCAACTGCTGAAGCGGCTCCACAGTGGATTAATGAGG 6240
DB 6181 TACTCAGAGCTCACTGTAAACCGAGCTCTGAGCGCACTGCATCAGTGAATAGCTCGG 6240
QY 6241 ACTGCTCTAGCCATGCTCCGGCTCGTGGCTTAAGGGATGTTTGGGATTTGGATATGCA 6300
DB 6241 AGTGATCCACTCCATGCTCCGGTTCCTGGCTAAGGAGACATCTGGGACTGGATATG 6300
QY 6301 TGTGACTGACTTCAAGACCTGGCTCCAGTCCAAACTCTGCGCGGGTTACCGGGAGTCC 6360
DB 6301 TGCTGAGCGACTTTAAGACCTGGCTGAAAGCGCAAGCTCATGCCCAAACTGCTGGGAT 6360
QY 6361 CTTTCTCTGTCATGCAAGCGGGGTACAAGGAGTCTGCGGGGGGACGGCATCATGCAAA 6420
DB 6361 CCTTGTGCTCTGCGACGCGGGTATAGGGGGTCTGCGGAGGAGACGGCATTAATGACA 6420
QY 6421 CCACCTGCCCATGCGGAGCAGATCGCCGGAACATGTCAAAAACGGTTTCCAATGAGGAT 6480
DB 6421 CTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACGATGAGGAT 6480
QY 6481 TAGGGCTTAGAACCTGCGAGCAACACGTGGGCAAGGAGCTTCCCATCAACGCAATACCA 6540
DB 6481 TCGGCTCTTAGGACCTGCGAGGAACATGTGGAGTGGGAGCTTCCCATTAACGCTTACAC 6540
QY 6541 CGGACCTTGCACACCTCCCGCGCCCAACTATTCCAGGGCGCTATGGCGGTGGCTG 6600
DB 6541 CGGCGCTCTGATCTCCCTCTGCTGCGCGCACTAATAGTTAGTTCGCTGTGGAGGTGTCT 6600
QY 6601 CTGAGGAGTACGTGAGGTTACGCTGTGGGGATTTCCACTACGTGAACGGGCAATGACCA 6660
DB 6601 CAGAGGAATACGTGGAGATAAGCGGGTGGGGGACTTCCACTAGTATCGGGTATGACTA 6660
QY 6661 CTGACAAAGTAAAGTGCCTATGCCAGGTTCCGGCCCCCGAATTTCTCAGCGAGGTGGATG 6720
DB 6661 CTGACAAATCTTAAATGCGCGTCCAGATCCCATCGCCCGAATTTTTCACAGAAATGGAG 6720
QY 6721 GAGTGGCTTGCACAGGTAGGCTCGCGGTGCAAACTCTTCTACGGGAGGACGCTCAGCT 6780
DB 6721 GGGTGGCTTACATAGGTTTGGCCCCCTTTCAGAGCCCTTCTGCGGGAGGAGGTATCAT 6780
QY 6781 TCCAGGCTCGGCTCAACCAATACTTGGTTCGGGTGCGAGCTCCCATTCGAGAGCCGCAAC 6840
DB 6781 TCAGAGTAGGACTCCACGAGTACCGGTGGGGTGCAGATTAACCTTTCGAGAGCCGCA 6840
QY 6841 ACGTAAAGTGTCTTACATTCATGCTCAGGATCCCTCCCACTTACAGCAGACAGCGGTA 6900
DB 6841 ACGTAGCGGTTCAGCGTCCATGCTCACTGATCCCTCCCACTAATAACAGCAGAGGGCG 6900
QY 6901 AGCTTAGGCTGCTAGAGGCTCCCGCTCTTTAGCCAGCTCATCAGCTAGCCAGTTGT 6960
DB 6901 GGAGAGGTTTGGCAGAGGGGTACCCCCCTTCTATGGCCAGCTCTCTGGCCAGCCAGCTGT 6960
QY 6961 CTGGCGCTTCTTGAAGGCGCATGCACTACCCACCACTGACTCCCGGACGCTGACCTCA 7020

Db 6961 CCGCTCCATCTCTCAAGGCAACTTGGACCGCCAAACCATGACTCCCTGAGCCGAGCTCA 7020
Qy 7021 TCAGGCGCAACCTCTTGTGGCGCAGAGATGGCGGAAACATCACTCTCGCGTGGAGTCAG 7080
Db 7021 TAGAGGCTAACCTCTCTGTGGAGCAGAGATGGCGGCAACATCAACAGGGTTGAGTCAG 7080
Qy 7081 AGAATAGGTAGTAATCTTGACTCTTTCGAAACCGCTTACCGCGGAGGGGATGAGAGG 7140
Db 7081 AGAACAAAGTGGTATCTTGACTCTCTTCGATCCGCTTGTGGCAGAGGAGATGAGCGGG 7140
Qy 7141 AGATATCCGTCGCGCGGAGATCCTCGCAAAATCCAGGAAGTTCCCTCAGCGCTTGCCCA 7200
Db 7141 AGTCTCCGTACCGCAGAAATCTCGGAAGTCTCGAGATTCGCCCGGCCCTCCCG 7200
Qy 7201 TATGGGCAACCGCCGAGCTCAATCTCTCACTAGTCTCGAGAGGACCGGACTAG 7260
Db 7201 TTTGGGCGCGCGGACTCAACACCCCGCTAGTAGAGAGCTGGAAAAAGCCTGACTAG 7260
Qy 7261 TCCCTCGGTGTACAGGATGCCATTCGACCTACCAAGGCTCCTCCAATACCACTC 7320
Db 7261 AACCACTGTGGTCCATGGCTGCCCGCTACCACTCCACGGTCCCTCTGTGCTCCGC 7320
Qy 7321 CACGAGAAAGAGACGGTTGTCTGACAGAAATCCAATGTGTCTTGCCTTGGCGGAGC 7380
Db 7321 CTGGAAAGAGCGTACGGTGTCTTCAACCAATCAACCTTACTCTGCTTGGCGGAGC 7380
Qy 7381 TCGCCACTAAGACTTTCGCTAGCTCCGATTCGTGGCGCTTGATAGCGGACCGGAGCG 7440
Db 7381 TTGCCACAAAGTTTGGCAGCTCTCAACTTCGGCAATACGGGCGACAATACGACAA 7440
Qy 7441 CCCTTCTGACTGGCTCCGACGAGCGTGCAAAAGATCCGAGCTTGGAGTCTGACTCCT 7500
Db 7441 CATCTCTGAGCGCGCCCTTCTGGCTGCCCGCCGACTCCGAGCTTGGAGTCTTATCTT 7500
Qy 7501 CCATGCCCTTGAAGGGGAGCGCGGACCCGATCTCAGCGACGGTCTTGGTCTA 7560
Db 7501 CCATGCCCTTGGAGGGAGCGCTGGGATCCGAGTCTCAGGACGGTCTAGTCTGA 7560
Qy 7561 CCGT---GAGTGAAGGAGCTAGTGAGGATGTCTGCTGCTCAATGCTCTATACGTGA 7617
Db 7561 CGGTCACTAGTGGGGCGACACGGAAGATGTCTGCTGCTCAATGCTCTATCTCTGA 7620
Qy 7618 CAGCGCCTGTATCACGCCATGGCTCGGGAGGAAGTAAGCTGCCCATCAACCCGTGA 7677
Db 7621 CAGCGCACTCGTCACCCCGTGGCTCGGGAAGAAACAAATCTGCCATCAACGCACTGA 7680
Qy 7678 GCAACTCTTGTGCTGCATCACCAACATGTCTACGCCCAACAACTCCCGAGCGCAAGCC 7737
Db 7681 GCAACTCTTGTGCTACGCCATCACAACTCTGGTGTATTCCACCCTTCAGCAGTCTGCC 7740
Qy 7738 TCCGGCAGAAAGGCTCACCTTTGACAGATTCGAAGTCTCTGGATGATCATTTACCGGAGC 7797
Db 7741 AAGGCAGAGAAAGTCACTTTGACAGACTGCAAGTCTTGGACAGCAATTTACCGAGC 7800
Qy 7798 TACTCAAGGAGATGAAGCGGAGCGGTCCACAGTTAAGGTAAAGCTTCTATCTATAGG 7857
Db 7801 TGCTCAAGGAGGTCAAAAGCAGCGCGTCAAAAGTGAAGGCTTAACTTGCTATCCGTAG 7860
Qy 7858 AGGCTCGAAGCTCAGCGCCCACTTCGGCCCAATCCAAATTTGGCTATGGGCAAAAG 7917
Db 7861 AAGCTTGCGCTCAGCGCCCACTTACGACCAATCCAAATTTGGCTATGGGCAAAAG 7920
Qy 7918 ACGTCCGGAACCTTATCAGAGAGGCGGTAAACCAACATCCGCTCGGTGGGAGGACTTGC 7977
Db 7921 ACGTCCGTGCTATGCCAGAAAGCGGTAGCGGCACATCACTCTCGTGTGGAAGACCTTC 7980
Qy 7978 TGAAGACACTGAACACCAATTGACACCAACCATATGGCAAAAGTGAAGTTTCTGCG 8037
Db 7981 TGAAGACAGTGTAAACCAATAGACACTACCACTATGCGCCCAAGAACAGGTTTCTGCG 8040
Qy 8038 TCCAAACAGAGAGGAGCGCGAAGCCAGCTCGCTTATCGTATTCACAGACTGGAG 8097
Db 8041 TTCAGCCTGAGAGGGGGTGTAAAGCCAGCTCGTCTCATCTGCTTCCCCACCTGGCG 8100

Qy 8098 TTCGTGTATGCGAGAAAGATGGCCCTTTTACGACGTGTCTTCCACCCCTTCTCAGGCCGTGA 8157
Db 8101 TGGCGGTGTGCGAGAAAGATGGCCCTGTACGACGTGTGTAGCAAGCTCCCCCTGGCGGTGA 8160
Qy 8158 TGGGCTCTCTCATACGATTTCAATACTCCCCCAAGCAGCGGGTCTGAGTCTCTGGTGAATA 8217
Db 8161 TGGGAGCTCTTACGATTTCCAATACTACACAGGACAGCGGGTTGAATCTCTCGTGAAG 8220
Qy 8218 CTTGGAATCAAGAAATGCCCTATGGGCTTCTCATATGACACCCCGCTGTTTGTACTCAA 8277
Db 8221 CTTGGAATCAAGAAAGACCCGATGGGGTCTCTGATATATCCCGCTGTTTGTACTCCA 8280
Qy 8278 CGTCACTCAGAGTGCATTTCTGTTGAGGAGTCAATTTTACCAATGTTGTGACTTGGCCC 8337
Db 8281 CAGTCACTCAGAGCGCATCTCGTACGAGAGGCAATTTTACCAATGTTGTGACTGGACC 8340
Qy 8338 CCGAGGCCAGACAGAGGCTCAAGGCTCTCACAGAGCGGCTTTACATCGGGGGTCCCTGA 8397
Db 8341 CCCAGCCCGGTGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTGGGGGCCCTCTTA 8400
Qy 8398 CTAACTCAAAAGGCGAGAACTGCGGTTATGCGCGGTGCGCGCAAGTGGGCTGCTGACGA 8457
Db 8401 CCAATTCAGAGGGGGAATACTGCGGTACCGCAGGTGCGCGCAGCGCTACTGACAA 8460
Qy 8458 CTAGCTCAGGACTCAGAGTGTCTGTAAACGAGGCGCACTGCGAGCCTGTCGAGCTGCAA 8517
Db 8461 CTAGCTGTGTGTAAACCCCTCACTGTCTACATCAAGGCCCGGCGAGCTGTCGAGCCGAG 8520
Qy 8518 AGCTCAGGACTCAGAGTGTCTGTAAACGAGGCGCACTGCGAGCCTGTCGAGCTGCAA 8517
Db 8521 GGCTCAGGACTCAGCATGCTCTGTGTGGCGGAGCTTAGTCTGTTATCTGTGAAAGTG 8580
Qy 8578 CCGGAACCCAGGAGATGCGGCGGCTTACGAGCTTTCAGGAGGCTATGACTAGTATT 8637
Db 8581 CCGGGGTCCAGGAGGACGCGCGGAGCTTGAGAGCTTTCAGGAGGCTATGACCAAGTACT 8640
Qy 8638 CCGCCCCCGGGGATCCGCCCCAACCAACGAGTACGAGCTTGGAGCTGATACATCATGTT 8697
Db 8641 CCGCCCCCGGGGACCCCCCAACAGAAATACGAGTACGAGCTTGGAGCTTATTAACATCATGCT 8700
Qy 8698 CTTCCAAATGTCTCAGTCTGCGCAGTGTCTTGGCAAAAGGCTATACCTCAACCCGTG 8757
Db 8701 CTTCCAAAGTGTCTGCTGCGCCACGACGCGCGCTGGAAGAGGCTTACTACTCTTACCCTG 8760
Qy 8758 ACCCCACCAACCCCTTTCAGCGGCTGCGTGGAGACAGCTAGACACACTTCCAATCACT 8817
Db 8761 ACCCTACAACCCCTCTCGCGAGAGCGCGTGGGAGACAGCAAGACACACTCCAGTCAAT 8820
Qy 8818 CTTGGCTAGGCAATATCATGTATGCGGCCCAACCTTATGGGCAAGGATGATCTGATGA 8877
Db 8821 CTTGGCTAGGCAATATCATGTATGCGGCCCAACCTTATGGGCAAGGATGATCTGATGA 8880
Qy 8878 CTCACCTTTTCTCATCTCTAGTCTCAAGAGCAACTTGAAGAGCCCTGATTTGTGCTGAGA 8937
Db 8881 CCAATTTCTTTAGGCTCTCATAGCAGGATCAGCTTGAACAGGCTTCTTAACTGTGAGA 8940
Qy 8938 TCTACGGGGTGTGCTACTCCATTTAGCCTTACCTTACCTCAGATCAATGAACGACTCC 8997
Db 8941 TCTACGGAGCTGCTACTCTCATAGAACCTTGGATCTACCTCCAATCAATCAAAAGACTCC 9000
Qy 8998 ATGCTCTTAGGCAATTTACACTCCAGATTTACTCTCAGGTGAGATCAATAGGTTGCTT 9057
Db 9001 ATGSCCTCAGCGCAATTTTCACTCCACAGTTTACTCTCCAGGTGAATCAATAGGTTGSCCG 9060
Qy 9058 CATGCTCAGGAACCTTGGGTTACCCCTTGGCACTCTGAGACATCTGGGCCAGAGAGTG 9117
Db 9061 CATGCTCAGAAACTTGGGGTCCCGCTTGGAGCTTGGAGACACCGGCCCCGGAGCG 9120
Qy 9118 TCCGCTTAAGCTACTCTCCAGGGGGGCGGCGACTTTGTGGCAGATACCTCTTTTA 9177
Db 9121 TCCGCTTAGGCTTCTCTCAGAGGAGCGAGGCTGCCATATGTGGCAAGTACCTCTTCA 9180

```
QY 9178 ACTGGCAGTAAGGACCAAGCTTAAACTCACTCAATCCCGCGCGCTCCAGCTGGACT 9237
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9181 ACTGGCAGTAAGAACAAAGCTCAAACTCACTCCAAATAGCGCGCTGGCGGCTGGACT 9240
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9238 TGTCTGGCTGGTTCTGCTGCTGTATACAGCGGGGAGACATATATACAGACCTGTCTGTTG 9297
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9241 TGTCCGGTTGGTTTCAAGGCTGGCTACAGCGGGGAGACATTTATACAGACGCTGTCTCATG 9300
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9298 CCGGACCCCGCTGGTTTCGGTTGTGCGCTACTCCTACTTCTGTAGGGGTAGGCATTTACC 9357
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9301 CCGGCGCCCGCTGGTTTCGGTTGTGCGCTACTCCTGCTGCGTGCAGGGGTAGGCATTTACC 9360
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9358 TGCTCCCCAACCGATGAACGGGAGCTAAACCACTCCAGGC-CTTAAGCCATTTCTGTTT 9416
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9361 TCCTCCCCAACCGATGAAGTTGGGTAAACACTCCGGCTCTTAGGCCATTTCTGTTT 9420
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9417 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9476
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9421 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9480
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9477 TT-----TCTTTTCCCT 9489
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9481 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9490 TCTTTAATGGTGGCTCCATCTTAGCCCTAGTGCACGGCTAGCTGTGAAGGTCCGTGAGCC 9549
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9541 TCTTTAATGGTGGCTCCATCTTAGCCCTAGTGCACGGCTAGCTGTGAAGGTCCGTGAGCC 9600
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9550 GCATGACTGCAGAGAGTGTGATGCTGATGCGCTCTTGCAGATCATGT 9595
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

QY 9601 GCATGACTGCAGAGAGTGTGATGCTGATGCGCTCTTGCAGATCATGT 9646
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

RESULT 15
US-09-995-937-1
; Sequence 1, Application US/09995937
; Publication No. US20030028010A1
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,937
; FILING DATE: 28-Nov-00 US20030028010A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,756
; FILING DATE: 04-May-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
;
;
; MOLECULE TYPE: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-937-1

Query Match 65.2%; Score 6254.8; DB 3; Length 9646;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 7558; Conservative 0; Mismatches 2037; Indels 51; Gaps 3;

QY 1 GCCAGCCCCCTGATGGGGGACACTCCACATGAATCACTCCCTCTGTAGGAATCACTG 60
Db 1 GCCAGCCCCCTGATGGGGGACACTCCACATGAATCACTCCCTCTGTAGGAATCACTG 60
QY 61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTGTAGTATGAGTGTCTGTCAGGCTCCAGAC 120
Db 61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTGTAGTATGAGTGTCTGTCAGGCTCCAGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATGGCAG 180
Db 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATGGCAG 180
QY 181 GACGACCGGGTCTCTTTCTTGGATCAACCCGCTCAATGCTTGAGATTTGGCGTGGCCCC 240
Db 181 GACGACCGGGTCTCTTTCTTGGATCAACCCGCTCAATGCTTGAGATTTGGCGTGGCCCC 240
QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAGGCCCTTGTGGTACTGCTGATAGG 300
Db 241 GCAAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAGGCCCTTGTGGTACTGCTGATAGG 300
QY 301 GTGCTTCGAGTGCCTCGGGAGGTCTCGTAGACCGTGCACCATCAGCACGAATCCTAAAC 360
Db 301 GTGCTTCGAGTGCCTCGGGAGGTCTCGTAGACCGTGCACCATCAGCACGAATCCTAAAC 360
QY 361 CTCAAGAGAAAACCAAAACGTAACCAACCGCCGCCACAGAGAGCTCAAGTTCCCGGGCG 420
Db 361 CTCAAGAGAAAACCAAAACGTAACCAACCGCCGCCACAGAGAGCTCAAGTTCCCGGGCG 420
QY 421 GTGCTCAGATGCTTGTGGTGGAGTTTACTGTTTCCGCGAGGGGCCCGAGTGGGTGTC 480
Db 421 GCGGTTCAGATGCTTGTGGTGGAGTTTACTGTTTCCGCGAGGGGCCCGAGTGGGTGTC 480
QY 481 GCGCGACTAGGAAGGCTTCGAGCGGTGCGAACCTCGTGGAGGCGCAACCTATCCCAA 540
Db 481 GCGCGAGAGGAAGACTTCGAGCGGTGCGAACCTCGAGGTAGACGTGAGCTATCCCAA 540
QY 541 AGGCTTCGCGACCCGAGGCGAGGCGCTGGGCTCAGCCCGGGTACCTTTGGCCCTCTATG 600
Db 541 AGGCACTGCGCCCGAGGCGAGGAGCTGGGCTCAGCCCGGGTACCTTTGGCCCTCTATG 600
QY 601 GCAATGAGGGCTGGGTGGGCGAGGATGGCTCTCTGTCACCCCGGGCTCCGGGCTAGTT 660
Db 601 GCAATGAGGGTGGGGGTGGGCGGATGGCTCTCTGTCCTCCCGGGCTCTCGGGCTAGCT 660
QY 661 GGGGCCCCACAGACCCCGGCTAGGTGCGTAACTTGGGTAAAGTCAATCGATACCCCTTA 720
Db 661 GGGGCCCCACAGACCCCGGCTAGGTGCGTAACTTGGGTAAAGTCAATCGATACCCCTTA 720
QY 721 CATCGGGCTTCGCGCATCTCATGGGGTACATTCGCTCGTGGGGCCCCCTTAGGGGCG 780
Db 721 CGTGGGGCTTCGCGCATCTCATGGGGTACATTCGCTCGTGGGGCCCCCTTAGGGGCG 780
QY 781 CTGCGAGGGCTTGGGACACACCGTGTCCGGGTCTTGGAGGACGGGTGAACATATGCAACAG 840
Db 781 CTGCGAGGGCTTGGGACACACCGTGTCCGGGTCTTGGAGGACGGGTGAACATATGCAACAG 840
QY 841 GGAACCTTCCCGGTGCTCTTTCTATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 841 GGAACCTTCCCGGTGCTCTTTCTATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 TCCAGCTTCCGCTTATGAAGTGGCGCAACGTGTCCGGGATATACCATGTCAGCAACGACT 960
Db 901 TCCAGCTTCCGCTTATGAAGTGGCGCAACGTGTCCGGGATATACCATGTCAGCAACGACT 960
```

Db 901 TGCCCGCTTCAAGCTACCAAGTGGCAATTCCTCGGGGCTTTACCAAGTCAACAATGATT 960
Qy 961 GTTCAAACTCAAGCATTTGTATGAGCGAGCGAGCTGATCATGCAATCTCCCGGTGG 1020
Db 961 GGCCTAACCTCGAGTATTTGTACGAGCGCGCGATGCCATCTTGCACTCCCGGTGG 1020
Qy 1021 TGCCCTGTGTTCAGGAGGTAAACAGCTCCCGTTGTCTGGGTAGCGCTCACTCCCAAGCTCG 1080
Db 1021 TCCCTTTCGTTCCGAGAGGTAAACGCTCCGAGGTGTTGGGTGGCGGTGACCCCAAGCTGG 1080
Qy 1081 CGGCCAGGAATGCCAGCGTCCCACTACGACAAATACGACCCACGCTGCACTGCTCGTTG 1140
Db 1081 CCACAGGAGCGCAAACTCCCAACAGCGAGCTTCGAGTCAATCGATGCTGCTTTGTCG 1140
Qy 1141 GGAGCGCTGTTCTGCTCGCTATGTAGTGGGGATCTCTCGGATCTTATTTCTCTCG 1200
Db 1141 GGAGCGCCACCTCTGCTCGCCCTCTACGTGGGGACCTGTGCGGTCTGCTCTTTCTTG 1200
Qy 1201 TCTCCAGCTGTTTCACTTCTCGCTCGCCGATGAGACAGTGCAGGATGCAACTGCT 1260
Db 1201 TTGGTCACTGTTTACCTTCTCTCCAGCGCCACTGGAGCGCAAGACTGCAATGTT 1260
Qy 1261 CAATCTATCCCGCCATGTATCAGGTCAACGATGCTTGGGATGATGATGAATGCT 1320
Db 1261 CTATCTATCCCGCCATATAAGGCTCATCGATGGCATGGGATATGATGAATGCT 1320
Qy 1321 CACCTACAAAGCCCTAGTGTGGGTGCGAGTGTCTCGGATCCCAAGCTGTCTGGACA 1380
Db 1321 CCCCTACGGAGCGTGTGGTGTAGCTCAGCTGCTCGGATCCCAAGCCATCATGGACA 1380
Qy 1381 TGCTGGGGGGCCACTGGGGAGTCTCGGGGCTTGGCTACTATTTCATGTTAGGGA 1440
Db 1381 TGATCGTGTGTCTACTGGGGAGTCTGGGGGATAGGATATTCTTCATGTTGGGA 1440
Qy 1441 ACTGGGCTAAGTGTCTGATTTGGGCTACTCTTTGGCGGTTGACGGGGAGACCCACA 1500
Db 1441 ACTGGGGAGGTCCTGTGTGTGTCTGCTATTGTCGGGTGCGACGCGGAAACCCACG 1500
Qy 1501 CGACGGGAGGTTGGCGGCGCACACCACTCCGGTTTCAAGTCCCTTTCTCATCTGGG 1560
Db 1501 TCACCGGGGAGTGGCGGCGCACACCACTGGCTGGCTTGTGGTCTCTTACACAGCG 1560
Qy 1561 CGTCTCAGAAAATCCAGCTTGTGATACCAAGCGAGCTGGCACATCAAGAGCTGCC 1620
Db 1561 CCAAGCAGAAATCCAACTGATCAACACCAACGCGAGTTGGCACATCAATAGCAGCGCT 1620
Qy 1621 TAAATTGCAATGCTCCTCCAACTGCTTTCTTTGGCGGCTGTTTACGACACAAAGT 1680
Db 1621 TGAATGCAATGAAGCTTTAAACCGGCTGGTTAGCAGGGCTCTTCTATCAGCACAAT 1680
Qy 1681 TCAACTCGTGGGTGGCGGAGCGCATGGCCAGCTGCGGCCCATTTGACTGTTTCGCC 1740
Db 1681 TCAACTCTTCAAGCTGTCTGAGAGGTGGCCAGCTGCGGAGCGCTTACCGATTTGCC 1740
Qy 1741 AGGGGTGGGGCCCATCACTATATAAGCTAAACAGCTCGATCGGATCAGAGGCTTATGCT 1800
Db 1741 AGGGCTGGGGTCTTATCAGTTATGCAACGGAAGCGGCTCGAAGACGCCCTACTGCT 1800
Qy 1801 GGCATTAAGCGGCTCGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
Db 1801 GGCATTAAGCGGCTCGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
Qy 1861 ATTGTTTCAACCGGCTGT 1920
Db 1861 ATTGTTTCAACCGGCTGT 1920
Qy 1921 ATAGCTGGGGAGGATGAGACAGAGCTGATGCTCTCAACAAACAGCGCTCCGCCACAAG 1980
Db 1921 ACAGCTGGGGTGCATAAGATAGGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
Qy 1981 GCAACTGTTGCGGTGTATCATGGATGAATAGTACTGGGTCACTAAGACGTCGGAGGTC 2040
Db 1981 GCAATTGTTGCGGTGTACCTGGATGAATCACTGGATTCACCAAGTGTGCGGAGCGC 2040

Qy 2041 CCCGCTGTAAATATCGGGGGGTGCGTAAACCGCACCTTTGATCTGCCCCACGAGCTGTTCC 2100
Db 2041 CCCCTTGTGTATCTCGAGGGGTGGGCAACACACCTTGTCTGCCCCACTGATTTGTTCC 2100
Qy 2101 GGAAGCACCCCGAGGTACTTTACAAAAATGTGGCTCGGGGCCCTGTTGACACCTAGGT 2160
Db 2101 GGAAGCATCCGGAAGCCACATATCTCTGGTGGGCTCCGGTCCCTGGATTAACACCGGT 2160
Qy 2161 GCTAGTAGACTACCCATACAGGCTTTGGCACTTACCCCTGCACTCTCAATTTTTTCCATCT 2220
Db 2161 GATGCTCACTACCCGTATAGGCTTTGGCACTATCTTTGTACCAATTAACACCATAT 2220
Qy 2221 TTAAGGTTTAGGATGTATGTGGGGGCGTGGAGCAGAGGCTCAATGGCGCATGAATGGA 2280
Db 2221 TCAAGTCAAGATGTACGTGGAGGGTGGAGCAGAGGCTGGAAGCGGCTGCAACTGGA 2280
Qy 2281 CTCGAGGAGCGCTGTAACTTTGGAGCAGGATAGGTTCAGAACTCAGCCCGCTGCTGC 2340
Db 2281 CGCGGGCGAAACGCTGTGATCTGGAAGACAGGACAGGTCAGAGCTCAGCCCATTTGCTGC 2340
Qy 2341 TGTCTAAACAGAGTGGCAGATACTGCCCTGTGCTTTTCAACACCTTACCGGCTTTATCCA 2400
Db 2341 TGTCCACCAACAGTGGAGGTCCTTCCGTGTTCTTTACGACCTTCCAGGCTTGTCCA 2400
Qy 2401 CTGGTTTGTATCCATCTCCATCAGAACTCGTGGACGTGCAATACCTGTACGGGTAGGGT 2460
Db 2401 CGGCTCTATCACCCTCCACCAGACATTTGTGGACGTGCACTTCTGACGGGTAGGGT 2460
Qy 2461 CAGGCTTGTCTCTTTGGCAATCAATGGAGTATACCTCTGTTGCTTTCTCTCTCTCTG 2520
Db 2461 CAAGCATCGGCTCTGGGCAATTAAGTGGAGTACGTCGTTCTCTCTCTCTCTCTCTG 2520
Qy 2521 CAGACGGCGGCTGTGCTGCTGTTGTGATGATGCTGCTGATAGCCAGCTCAGAGCG 2580
Db 2521 CAGACGGCGGCTGTGCTGCTGTTGTGATGATGCTGCTGATAGCCAGCTCAGAGCG 2580
Qy 2581 CTTAGAGAACTTGGTGTCTCAATCGCGGCTCGTGGCGGAGCGCATGTTCTCT 2640
Db 2581 CTTTGGAGAACTCTGATATCTCAATGACAGATCTCTGCGGAGCGACGACGGTCTGTGT 2640
Qy 2641 CTTTCTTGTGTTCTCTGCGCGCTCTGTTACATTAAGGCGAGCTGCTCTGCGGCGG 2700
Db 2641 CTTTCTGTTGTTCTCTGCTTGTGGTGTATCTGAAGGTTAGTGGTGGTGGTGGTGGTGG 2700
Qy 2701 CGTATGCTTTTATGCGGTATGCGCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Db 2701 TCTACGCTTCTACCGGATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Qy 2761 CTTACGCTTGGACCGGAGATGGCTGATCGTGGGGGTGGGTTCTTGTAGTGTCTGG 2820
Db 2761 CATAAGCACTGGACACAGGAGTGGCGGCTGCTGTTGGGGGTGTTCTTGTGGGTAA 2820
Qy 2821 TATTCTTGTGCTGCTACCATCTACAAAGTGTCTCTCACTAGGCTCATATGTTGGTTAC 2880
Db 2821 TGGCGCTGATCTCTGCGCATTTTAAAGCGCTACATCAGCTGTTGATGTTGGTCTTC 2880
Qy 2881 AATACTTTTATCAGCAGAGCGAGCGACATGCAAGTGTGGGTCCCCCTCTCAACGTTT 2940
Db 2881 AGTATTTTCTGACAGAGTAGAAGCGAACTGACAGTGTGGTTCCTCCCTCAACGTTT 2940
Qy 2941 GGGAGGCGCGATGCCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3000
Db 2941 GGGGGGGCGGATGCGCTCATCTTACTCATGTTGTTGTACACCGGCTCTGTTGTTTGT 3000
Qy 3001 ACATCACAACCTCTGCTCGCATATCTGCGGCTCATCTGTTCCAGCTGTCATAA 3060
Db 3001 ACATCACAACCTCTCTGCGCATCTTCTGCGGCTCTTCTGAGTCTTCAAGCGGTTTGC 3060
Qy 3061 CGAGAGTGGCTGCTCTGTCGCTCAAGGGCTCATCTGTCGATGCTGTTAGTGGGAA 3120
Db 3061 TTAAGTCT 3120

QY 3121 AAGTCGCGGGGTCAATTATGTGTCAAATGGTCTTCATGAAGCTGGCGCGGTGACAGGTA 3180
Db AGATAGCGGGAGTCAATTACGTGCAAAATGGCCATCATCAAGTTAGGGCGGTACTGGCA 3180
QY 3181 CGTAGCTTTATACCATCTTACCCACACTGCGGACTGGGCCACGCGGCCCTACAGACC 3240
Db CCTATGTGTATAACCATCTCACCCCTCTTCGAGACTGGCGGCACAAACGGCCCTGCGAGATC 3240
QY 3241 TTGCGGTGGCGGTAGAGCCCGTCTCTTCGCGCATGGAGACCAAGGTTCATCACTGGG 3300
Db TGGCCGTGGCTGTGAAACCAAGTCGTCTCTCCGAAATGGAGACCAAGCTCATCAGTGGG 3300
QY 3301 GAGCAGACACCGCTCGGTGTGGGACATCATCTTGGGTCTACCGTCTCCGCGCCGAAAGG 3360
Db GGGCAGATACCGCGGTGGGTGACATCATCAACGGCTTGGCCGTCTCGCCCGTAGGG 3360
QY 3361 GGAAGGATATTTTGGGACCGGTGATGCTCTGAAAGGCAAGGTGGGACTCCCTTG 3420
Db GCCAGGATATCTGTTGGGCCAGCGACCGAAGTGGTCTCCAAAGGGTGGAGTTGCTGG 3420
QY 3421 CGCCCATCAGCGCTACTCCAAACAAACGCGGGCGGTACTTGGTTCATCATCACTAGCC 3480
Db CGCCCATCAGCGGTACGCCAGCAGACGAGAGGCTCTTAGGGTGTATATCACCAAGCC 3480
QY 3481 TCACAGCGCGGACAAAGAACCAAGTCCGAAAGGGAGTTCAAGTGGTTTCTACCGCAACAC 3540
Db TGACTGGCGCGGACAAACCAAGTGGAGGTGAGTCCAGATCGTGTCAACTGCTACCC 3540
QY 3541 AATCTTTCTGGCGACTGTCATCAACGGGTGTGCTGACTGTCTACCATGGCGTGGCT 3600
Db AAACCTTCTCGGCAACGTGCATCAATGGGGTATGTGACTGTCTACCAAGGGCGCGAA 3600
QY 3601 CGAAGACCTAGCGGTCCAAAGGTCCAAACCAAGTACCAATGACCAATGAGACCTGG 3660
Db CGAGACCATCGCATCACCAAGGTCTCTGTCATCGAGTGTATACCAATGAGACCAAG 3660
QY 3661 ACCTCGTGGTGGAGGCGCCCGCGGGCGCTCCATGACACCATGACAGCTGTGCA 3720
Db ACCTTGTGGGTGGCGCGCTCTCAAGGTTCGCGTCTATTTGACACCCCTGCACTCGGCT 3720
QY 3721 GCTCGGACTTTACTTGTGTCAGACATGCTGATGTCTATCCGTTGGCGCGCGGAGGG 3780
Db CCTCGGACTTTACTTGTGTCAGAGGACGCGGATGTCTATCCGTTGGCGCGCGGAGGTG 3780
QY 3781 ACAGCAGGGAAGTCTACTCTCCCGCGCGCTCTCTTACCTGAAAGGCTCTCGGGT 3840
Db ATAGCAGGGTAGCTGCTTTGCGCCCGCGCCATTTCTACTTGAAGGCTCTCGGGG 3840
QY 3841 GTCCATTGCTTTGCCCTTGGGGCAGCTGCGGGCGTCTTCCGGGCTGCTGTGTGCAACC 3900
Db GTCCGCTGTTGTGCGCGGACACGCGTGGGCTATTTCAGGGCGCGGCTGTGCAACC 3900
QY 3901 GGGGGTCCGAAGGGGTGACTTCATACCGTTGAGTCTATGAAATACCATGCGGT 3960
Db GTGAGTGGCTTAAGCGGTGGACTTTATCTCTGTGGAGAACCTTAGAGACAACCATGAGAT 3960
QY 3961 CTCCTGCTTACAGACAACCTCAACCCCGCGCTGTACCGCAGACATTTCCAAGTGGAC 4020
Db CCGCGGTGTTACGAGACAACCTCTCTCCACAGAGTGGCCCGAGAGTCTTCAGGTGGCC 4020
QY 4021 ATCTGACGCTCTACTGCGAGCGGCAAGACCAAAAGTGGCGGTGCGTATGAGGCC 4080
Db ACCTGATGCTCCACCGCGCAGGTAAGACCAACAAAGTCCCGGCTGCGTACGAGGCC 4080
QY 4081 AAGGTACAGGTGCTGTCGTAACCGCTCGTTGCGCGCACTTAGGGTTGGGGGT 4140
Db AGGGTACAAAGTGTGGTGTCTCAACCCCTCTGTTGCTGCAACGCTGGGCTTTGGTGT 4140
QY 4141 ATATGTCCAAGGCACACGCTATCGACCTAAACATCAGAACTGGGTAAAGACCATTACCA 4200
Db ACATGTCCAGGCCCATGGGGTGTATCTCTAAATATCAGGACCGGGTGAACAATTACCA 4200
QY 4201 CGGGCGGCTCCATTAGTACTCCACCTATGGCAAGTTCTTCCGACGGTGGCTGTCTG 4260

Db 4201 CTGGCAGCCCATCAGCTACTCCACCTACGGCAAGTTCTTCCGACGCGGGTCTCAG 4260
QY 4261 GGGCGCTATGACATCATATAATATGTATGATGTCACCTCAACTGACTCGACTACCATCT 4320
Db GAGGTGCTTATGACATATAATATTTGTACGAGTGGCACTCCACGATGCAATCACT 4320
QY 4321 TGGGCATCGGCACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGCGGTCTGTCGTCG 4380
Db TGGGCATCGGCACAGTCTCTTACCAAGCAGAGACTGCGGGGCGAGACTGGTTGTGTCG 4380
QY 4381 CCACGGCTACACTTCGGGATCGGTTACCGTGCACACACCCCAATATCGAGGAAATAGGCC 4440
Db CCACTGCTACCCCTCCGGGCTCCGTCACCTGTCTCCATCTCAACATCGAGGAGTGTCTC 4440
QY 4441 TGTCCAAACAATGGAGAGATCCCTTCTATGCAAAAGCCATCCCATTTGAGGCCATCAAG 4500
Db TGTCCACACCGAGAGATCCCTTTTACGGAAGGCTATCCCTCGAGGTGATCAAG 4500
QY 4501 GGGGAGGACATCTCATTTTCTGCCATTCACAGAGAAATGTGACGAGCTCGCGCAAGC 4560
Db GGGGAAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGGACGAGCTCGCGCGAAGC 4560
QY 4561 TGACAGGCTCGACTGAACTGATGATATTTACCGGGCTTGTATGTCCGTCTATC 4620
Db TGGTGCATTTGGGCATCAATGCGGTGGCTACTACCGCGTCTTCCAGTGTCTGTCTC 4620
QY 4621 CGCCTATCGGAGAGCTGCTGTCGTGGCAACAGACGCTCTAATGACGGGTTTCACCGGG 4680
Db CGACAGCGGGGATGTTGTCGTGTGTCGACCGATGCTCTATGACTGGCTTTACCGGG 4680
QY 4681 ATTTGACTCAGTGTGACTGCAATATATGTGTCCACCCAGACAGTTCAGCTTGG 4740
Db ACTTCGACTCTGTGATGACTGCAACAGTGTGTCTCAGTGTGATTTTTCAGCTTGG 4740
QY 4741 ATCCACCTTCCATNTGAGACGACCGTCCGCCAAAGACGCGGTGTGCGCTCGCAAC 4800
Db ACCCTACCTTTACCAATTTGAGACAACACGCTCCCCAGGATGCTCTCTCCAGGACTCAAC 4800
QY 4801 GCGAGGTGAGAACTTGGCAGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC 4860
Db GCGGGGAGGACTTGGCAGGGGGAAGCAGGCACTTACAGATTTGTGGCACCGGGGAGC 4860
QY 4861 GCGCTCGGGCATGTTTCAATTTCTTGGTCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
Db GCGCTCGGGCATGTTTCACTCGTCCGTCTCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
QY 4921 GGTATGAGCTCAGCGCCCTGAGACCTCGTTAGTGTGGGCTTACCTAAATACACAG 4980
Db GGTATGAGCTCAGCGCCCGCAGAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCG 4980
QY 4981 GGTGCGGCTTCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTCAGGSCCTCACCC 5040
Db GGTGCGGCTTCCAGGACCATCTTGAATTTTGGGAGGCGCTTTTTCAGGSCCTCACCTC 5040
QY 5041 ACATAGATGCCACTTCTCTGCCAGACTTAAACAGCAGGAGACAACTTCTTCTTACCTGG 5100
Db ATATAGATGCCACTTCTATCCAGACAAAGCAGTGGGAGAACTTCTTCTTACCTGG 5100
QY 5101 TGGCATATCAAGCTACAGTGTGCGCAGGGTCAAAGTCCACCTCATCTGTGGGACCAAA 5160
Db TAGCGTACAAAGCCACCGTGTGCGTAGGGTCAAAGCCCTCCCGCATCGTGGGACCA 5160
QY 5161 TGTGAAAGTGTCTCATCGGCTGAAACCTTACATCGACGGGCCAACACCCCTGCTGTATA 5220
Db TGTGAAAGTGTGTGATCGGCTTAAACCCACCTTACATGGGGCAACACCCCTGCTGTATA 5220
QY 5221 GCTAGGAGCGCTCAAAATGAGGTCTCTTCCACACACCCCAATAAATACATCATGG 5280
Db GACTGGGCGCTTTCAGATGAGTCACTCCCTGACGCAACCAATCACCATAATCATCA 5280
QY 5281 CATGATGTGGGCTGACTGGAGGTGTCTCATGTAGCACTGGGTGTGTGTAGCGGAGTCC 5340

5281 CATGCAATGTCGGCGCACTTGGAGGTGCTCAGAGCACCTGGGTGCTGCTTGGCGGCGTCC 5340
5341 TTGCAGCTTTGGCGCATCTGCTCAGACAGCAGAGTGTGGTCACTTGTGGCGAGGATCA 5400
5341 TGGCTGCTCTGGCGCGGTATTGCTCTCAACAGGCTTGGTGTATAGTGGCGAGATTG 5400
5401 TCTTGTCCGGGAAGCAGCTGTCTGCTCCGACAGGGAAGTCTCTACACAGGAGTTCGATG 5460
5401 TCTTGTCCGGGAAGCAGGCAATATATCTGACAGAGGAGTCTCTACACAGGAGTTCGATG 5460
5461 AGATGGAAGTGTGCTCACAACCTTCTTACATCGACAGGGAATGACGTCGCGGAGC 5520
5461 AGATGGAAGTGTGCTCACAACCTTCTTACATCGACAGGGAATGATGCTCGCTGAGC 5520
5521 AATTCAAGCAAAAGGCGCTCGGCTTGTGCAACGGCCCAACAGCGAGCGAGGCTGCTG 5580
5521 AGTTCAAGCAAGAGGCGCTCGGCTTCTGACAGACCGGCTCCCGCAAGCAGAGGTTATCA 5580
5581 CTCCTGCTGAGTCCAACTGGCGAGCCCTTGAGACCTTCTGGGCGAAGCAATGTGGA 5640
5581 CCCCTGCTGACAGCAACTGGCAGAACTCGAGGCTTCTTGGGCGAAGCACAATGTGA 5640
5641 ATTTCAACAGGGAATACAGTACTAGCAGGCTTATCCACTCTGCTGGGAAACCCCGCA 5700
5641 ATTTCAACAGGGAATACAGTACTAGCAGGCTTATCCACTCTGCTGGGAAACCCCGCA 5700
5701 TAGCATCATGATGGCAATTTACAGCTTCTATCACTAGCCGCTCACCACCAAAACACC 5760
5701 TTGCTTCAATGATGGCTTTTACAGCTGCGGTGCGGCGCTGTCACGCTGCTGTAACCCGCA 5760
5761 TCCTGTTTAAACATCTTGGGGGAGTGGGTGCTGCCAACTCGCTCTCCAGCGGCTGCT 5820
5761 TCCTTCAACATATTGGGGGAGTGGGTGCTGCCAGCTCGCGCGCCCGGTCGCGCTA 5820
5821 CAGCTTTGCGGGCGCGGATCGCCGAGCGGCTGTGGCAGCATAGGCTTTGGGAGG 5880
5821 CCGCTTTGCGGGCGCTGCTGCTAGCTGCGCGCGCATCGCGAGCTTTGGAATGGGGAAG 5880
5881 TGCTCGTGGACATCTTGGCGGCTATGGGGCAGGGTAGCGGGCACTGTGGCTTTTA 5940
5881 TCCTCGTGGACATCTTGGCGGCTATGGGGCAGGGTAGCGGGCACTGTGGCTTTTA 5940
5941 AGTCAATGAGCGGAGTGGCTTCCACCGAGGACCTGGTCAACTTCTCCCTGCCATCC 6000
5941 AGATCAATGAGCGGTGAGTCCCTTCCACGAGGACCTGGTCAACTGCTGCGCGCATCC 6000
6001 TCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6060
6001 TCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6060
6061 GCGCGGAGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCTTGGCTTGGCGGGTA 6120
6061 GCGCGGAGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCTTGGCTTGGCGGGTA 6120
6121 ACCAGCTCTCCCTACGCACTATGCTGCTGAGAGCGCTGACAGCTGCTCACTCAGA 6180
6121 ACCAGCTCTCCCTACGCACTATGCTGCTGAGAGCGCTGACAGCTGCTCACTCAGA 6180
6181 TCCTCTCTAGCTTACCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 6240
6181 TACTCAGCAGCTCACTGTAACCCAGCTCTGAGCGAGCTGATCAGTGGGATAGCTCGG 6240
6241 ACTGCTCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6300
6241 AGTGACCACTTCAATGCTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6300
6301 TGTTCAGCTTCAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6360
6301 TGCTGAGGAGCTTTAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6360
6361 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6420
6361 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6420

6421 CCACCTGCTCCCATTCGGAGCAGAGATCGCGGACATGTCAAAAACGGTTCCATGAGGATCG 6480
6421 CTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCG 6480
6481 TAGGGCTTAGAACCTTGCAGCAACAGTGGAACAGTTCGCCCATCAACAGCATACACCA 6540
6481 TGGTCTTAGGACCTTGCAGCAACAGTGGAACAGTTCGCCCATCAACAGCATACACCA 6540
6541 CCGGACCTTTGCACACCTCCCGGCGCCCAACTATTCAGGGGCGCTATGGCGGCTGCTG 6600
6541 CCGGCGCTGTACTTCCCTTCTCGCGCCGAACCTATAAGTTTCGCGCTGTGGAGGCTGCTG 6600
6601 CTGAGGAGTACGTGGAGTTAGCGGTGCGGGGATTTCCACTACGTAGCAGGCGATACCA 6660
6601 CAGAGGAATACGTGGAGATAAGCGGGTGGGGACTTCCACTACGTATCGGGTATGACTA 6660
6661 CTGACAACTGAAAGTCCCATGCCAGTTCCGGCGCCCGAACTTCTTACGGAGGTGGATG 6720
6661 CTGACAACTTAAATGCCCGTGCAGATCCCATCGCCGAACTTTTTCACAGAATTGGACG 6720
6721 GAGTGGGTTGCACAGGTACGCTTCGGCGTGCAGAACTTCTTACCGGAGGAGCTCACGT 6780
6721 GGGTGGCTTACATAGTTTGGCGCCCTTGCAGCCCTTGTGCGGGAGGAGTATCAT 6780
6781 TCAGGTGGGCTCAACCAATATCTTGGTGGGTGCGAGCTCCCATCGGAGCCCGACCGG 6840
6781 TCAGGTGGGCTCAACCAATATCTTGGTGGGTGCGAGTTCCTTACCTTGGAGCCCGAACCG 6840
6841 ACCTACAGTCTTACTTCCATGCTCAGCGATCCCTCCCACTTACAGCAGAGACGGCTA 6900
6841 ACCTAGCGGTGAGTCTTCCATGCTCAGTATCCCTCCCACTTACAGCAGAGGCGCGG 6900
6901 AGCTAGGCTGGCTAGAGGCTTCCCGCTTCTTACCGAGCTCATAGCTAGGCGAGTGT 6960
6901 GAGAGGTTGGCGAGAGGCTCACCCCTTCTATGGCCAGCTTCTCGGCCAGCGAGTGT 6960
6961 CTGCGCTTCTTTGAGGCGACATGACACTACCACTGACTCCCGGAGCGCTGACCTCA 7020
6961 CCGCTCCATCTCTCAAGGCAACTTGCACCGCCCAACCATGATCCCTTGCAGCGAGTCA 7020
7021 TCAGGCGCACTTCTTGGCGGAGAGATGGCGGAAACATCACCTCGGTGGAGTCAG 7080
7021 TAGAGGTAACCTTCTGAGGAGAGATGGCGGAAACATCACAGGTTGAGTCAG 7080
7081 AGAATAAGTAGTAATTTCTGAGCTTCTTGGAAACCGCTTTCAGCGGAGGAGGATGAGGG 7140
7081 AGAATAAGTAGTAATTTCTGAGCTTCTTGGAAACCGCTTTCAGCGGAGGAGGATGAGGG 7140
7141 AGATATCCGTGCGCGGAGATCTGCGAAATCCAGAAATTCAGGAGTTCCTCTCAGCGTTGCCA 7200
7141 AGGTCTCGGTACCGCAGAAATCTCGGAAATCTCGGAGATTCGCGCGGCGCTGCGCG 7200
7201 TAGGGCAGCGCGGAGTCAATCTCTCAGTCTGCTGAGTCTTGGAGGACCGGAGTACG 7260
7201 TTTGGGCGCGCGGAGTCAATCTCTCAGTCTGCTGAGTCTTGGAGGAGTGGAAAGGCTGACTACG 7260
7261 TCCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7320
7261 AACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7320
7321 CACGAGAAAGAGGAGTGTCTCTGAGCAATCCAATGTGTCTTCTGCTTGGCGGAGC 7380
7321 CTGAGAAAGAGGAGTGTCTCTGAGCAATCCAATGTGTCTTCTGCTTGGCGGAGC 7380
7381 TCGCCATTAAGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7440
7381 TTGCCACCAAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7440
7441 CCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7500
7441 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7500

Qy	7501	CCATGCCCCCCCTTTGAAGGGGAGCCGGGGAGACCCCGATCTCACGACGCGGTCTTGTGCTA	7560
Db	7501	CCATGCCCCCCCTTGGAGGGGGAGCCTTGGGATCCGGATCTCTCAGCGACGGGTCAATGTCGA	7560
Qy	7561	CCGT- --GAGTGGAGGCGCTAGTGAGGATCTGCTCTGCTGCTCAATGTCCTATACCTGGA	7617
Db	7561	CGGTCACTAGTGGGGCCGACACGGAAGATGTCGTGTGCTGCTCAATGTCCTATTCCTCGA	7620
Qy	7618	CAGCGCCCTTGATCACGCCATCGCCTCGGAGGAAAGTAAGCTGCCCATCAACCCCTTGA	7677
Db	7621	CAGCGCACTCGTCAACCCGTCGCTCGGAGGAGAACAAAACCTGCCCATCAACGCACTGA	7680
Qy	7678	GCAACTCTTTGCTGCGTCAACAACAATGTTACGCCACAACAATCCCGCAGCGCAAGCC	7737
Db	7681	GCAACTCGTTGCTACGCCCATACAATCTGGTGATTCCACCACTTCACGCAGTGTCTGCC	7740
Qy	7738	TCCGGCAGAGAGGTCACCTTTGACAGATTGCNAGTCCCTGGATGATCATATCCGGGACG	7797
Db	7741	AAAGGCAGAGAAAGTCACATTTGACAGACTGCAAGTCTTGGACAGCAATTCACGAGCG	7800
Qy	7798	TACTCAAGGAGATCAAGCGAAGCGCTCCACAGTTAAGGCTAAAGTCTTATCTATATAGAGG	7857
Db	7801	TGCTCAGGAGGTCAAAGCAGCGCGCTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGG	7860
Qy	7858	AGGCTGCAAGCTGACGCCGCCACATTCGGCCAAATCCAAATTTGGCTATGGGGCAAAAG	7917
Db	7861	AAGCTTGACGCTGACGCCGCCACATTCAGCAAAATCCAAAGTTTGGCTATGGGGCAAAAG	7920
Qy	7918	ACGTCCGGAACCTATCAGCAGGGCCGTTTAACACATCCGCTCGGTGGAGGACTTGC	7977
Db	7921	ACGTCCGTGTGCATTCGCAGAAAGCCGTAGCCCAATCAACTCCGTGTGGAAAGACCTTC	7980
Qy	7978	TGGAAGACACTGAAACCAATTCACACACCATCATGGCAAAAGTAGAGTTTCTGCG	8037
Db	7981	TGGAAGACAGTGTAAACCAATAGACACTACCATCATGGCCAGAACGAGGTTTTCTGCG	8040
Qy	8038	TCCAAACAGAGAGGAGGCGCAAGCCAGCTCGCCTTATCGTATTTCCACAGACTGGAG	8097
Db	8041	TTCAGCCTTGAGAGGGGGTCTGAAGCCAGCTCGTCTCATCTGTTTCCCGACCTGGCG	8100
Qy	8098	TTCTGTGTATGCGAGAGANTGCCCTTTACAGAGTGGTCTCACCCCTTCCTCAGGCCGTGA	8157
Db	8101	TGCGCGTGTGCGAGAAATGCGCCCTTACGACGTGGTTAGCAAGCTCCCCCTCGCCGTGA	8160
Qy	8158	TGGCTCCTCATACGGAATTTCAATACTCCCCCAAGCAGCGGTCGAGTCTCTGTGAATA	8217
Db	8161	TGGGAAGTCTCTACGGATTCCAATACTACCGAGACAGCGGTTGAATTTCTCGTGCAG	8220
Qy	8218	CCTGGAATCAAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGACTCAA	8277
Db	8221	CGTGGAGTCCAAGAGACCCCGATGGGTTCTCGTATGATACCCGCTGTTTGACTCCA	8280
Qy	8278	CGGTCACTGAGAGTGACANTTCGTGTGAGGAGTCAATTTACAAATGTTGACTTGGCCC	8337
Db	8281	CAGTCACTGAGAGGCATCCGTATCGGAGGAGGCAATTTACCAATGTTTGACCTGGACC	8340
Qy	8338	CCGAGGCCAGACAGCCCATAGTTCGCTACAGAGCGGCTTATCATCGGGGTCCTCCCTGA	8397
Db	8341	CCCAAGCCCGGTGGCCCATCAAGTCCCTTCACTGAGAGGCTTTATGTTGGGGGGCCCTCTTA	8400
Qy	8398	CTAACTCAAAAGGGCAGAACTGCCGTTATCCCGGTCGCGCAGAGTGGGCTGTGACGA	8457
Db	8401	CCAAATCAAGGGGGGAAACTTGGCGTACCGCAGGTGCGCGCAGCGGGGTACTGACAA	8460
Qy	8458	CTAGTCGGTAAATACCTCATATGTTACTTTGAAGGCCACTTGCAGCCTGTGAGCTGCAA	8517
Db	8461	CTAGCTGTGGTAAACACCCCTCACTTGTCTACATCAAGAGGCCGGGCGAGCCTGTGAGCCGAG	8520
Qy	8518	AGCTCCAGGACTGCACGATGCTCGTGAACGAGACGACCTTGTGTTATCTGTGAAGCG	8577
Db	8521	GGCTCCAGGACTGCACCAATGCTCGGTGTGGCGACGACTTAGTCTGTTATCTGTGAAAGTG	8580
Qy	8578	CGGAAACCCAGGAGGATGCGCGGGCCCTTACGAGCGCTTACCGAGGCGCTATGACTAGGTATT	8637

D	b	8581	CGGGGTCCAGGAGGACGGCGGAGCCTGAGAGCCTTACGGAGGCTATGACCAGGTACT	8640
Q	y	8638	CCGCCCCCCCGGGGATCGCCCCCAACAGAAATACGACCTGGAGCTTGATAAATCATGT	8697
D	b	8641	CCGCCCCCCCGGGGACCCCCACAACAGAAATACGACTTGGAGCTTTAAACATCATGCT	8700
Q	y	8698	CCTCCAATGTGCAGTCGGGACGATGTGCATCTGGCAAAGGGTATACTACCTCACCCGTG	8757
D	b	8701	CCTCCAAGTGTGAGTCGCCACGACGGGGCTGGAAGAGGGTCTACTACTTACCOCGTG	8760
Q	y	8758	ACCCACCAACCCCTTGACGGGCTCGCTGGGAGACGCTAGACACATCCCAATCAACT	8817
D	b	8761	ACCTCAACCCCTCGGAGAGCCGCTGGGAGACAGCAACACATCTCCAGTCAATT	8820
Q	y	8818	CTTGGCTAGGCAATATCATCATGTATGCGCCACCTATGGGCAAGGATGATTCTGATGA	8877
D	b	8821	CCTGGCTAGGCAATAATCATGTGTTGCCCCACACTGTGGGCGAGATGATCTGATGA	8880
Q	y	8878	CTCACTTTTTCTCCATCCTTCTAGCTCAAGAGCAAATTGAAAAGCCCTGGATGTCAGA	8937
D	b	8881	CCCATTTCTTTTAGCGTCTCATAGCCAGGATCAGCTTGAAACAGGCTCTTAACGTGAGA	8940
Q	y	8938	TCTACGGGCTTGTACTCCATTGAGCCATTGACCTACTCAGATCTTGAACGACTCC	8997
D	b	8941	TCTACGGAGCTGTACTCCATAGAACCACTGGATCTACTCCAACTATTCAAAGACTCC	9000
Q	y	8998	ATGCTCTTAGCGATTTACACTCCACAGTTACTCTCCAGTGAGATCAATAGGGTGGCTT	9057
D	b	9001	ATGGCTCAGCGATTTTCACTCCAAGTTACTCTCCAGTGAAATCAATAGGGTGGCCG	9060
Q	y	9058	CATGCTTCAGGAAACTTTGGGGTACCACCTTGGCAACTGGAGACATCGGGCCAGAGTG	9117
D	b	9061	CATGCTTCAGAAACTTTGGGGTCCCGCCCTTGGAGCTTGGAGACACCGGGCCGGAGCG	9120
Q	y	9118	TCCGCGTAACTACTGTCCAGGGGGAGGGCCGCCACTTGTGGCAGATACCTCTTTA	9177
D	b	9121	TCCGCGTAGGCTTCTGTCCAGAGGAGCGAGGCTGCCATATGTGGCAAGTACTCTTCA	9180
Q	y	9178	ACTGGGAGTTAAGGACCAAGCTTAAACTACTCCAATCCCGCCGGGTCCAGCTGGACT	9237
D	b	9181	ACTGGGAGTTAAGAACAAAGCTCAAACACTCAATAGCGGCCGTGGCCGGCTGGACT	9240
Q	y	9238	TGCTGCGCTGGTTCGTCGTGTTACAGCGGGGAGACATATACAGCCTGTCTCGTG	9297
D	b	9241	TGTCCGGTTGGTTCAACGGCTGGCTACAGCGGGGAGACATTTATCAGCGGTGTCTCATG	9300
Q	y	9298	CCGACCCCCTGGTTCCGCTGTGCTACTCTCTTCTGTAGGGGTAGGCATTTACC	9357
D	b	9301	CCCGCCCCCGTGTGTGTTTGGCTACTCTCTGCTCGCTGCGAGGGTAGGCATCTACC	9360
Q	y	9358	TGCTCCCCAAACCGATGAACGGGAGCTAAACCACTCCAGGC-CTTAAGGCCATTTCCGTGTT	9416
D	b	9361	TCCTCCCAACCGATGAAGGTTGGGGTAACAATCCGGCCCTCTTAGGCCATTTCCGTGTT	9420
Q	y	9417	TT	9476
D	b	9421	TT	9480
Q	y	9477	TT-----TCTTTTTCCCT	9489
D	b	9481	TTTTTTCCTTTTTTTTTTTTTTTTTTTTTTTTTTCTTCTCTTTTTTCTTTTCTCTTCTCT	9540
Q	y	9490	TCTTTAATGGTGGTCCCATCTTAGCCCTTAGTFCACGGCTAGCTGTGAAGGTCGGTGAGCC	9549
D	b	9541	TCTTTAATGGTGGTCCCATCTTAGCCCTTAGTFCACGGCTAGCTGTGAAGGTCGGTGAGCC	9600
Q	y	9550	GCACTGACGACAGAGTGTGATCTAGGCTCTCTGCAGATCATGT	9595
D	b	9601	GCACTGACGACAGAGTGTGATCTAGGCTCTCTGCAGATCATGT	9646

us-09-662-454-4.rnpbm

Mon Dec 12 10:04:26 2005

Job time : 4536 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 20:15:48 ; Search time 509 Seconds
(without alignments)
7047.417 Million cell updates/sec

Title: US-09-662-454-4
Perfect score: 9595
Sequence: 1 gccagccctgatggggc.....ggcctctctgcagatcatgt 9595

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New.*
1: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	6335.8	66.0	9599	6	US-10-985-205-1
2	6138.4	64.0	7979	6	US-10-509-921-11
3	6136.8	64.0	7979	6	US-10-509-921-12
4	5905.6	61.5	7989	6	US-10-509-921-8
5	5733.6	59.8	7979	6	US-10-509-921-9
6	5732	59.7	7979	6	US-10-509-921-10
7	5516.8	57.5	7989	6	US-10-509-921-13
8	5318.4	55.4	7989	7	US-11-119-330-1
9	5318.4	55.4	7992	7	US-11-111-686-1
10	5316.8	55.4	7992	7	US-11-111-686-6
11	5315.2	55.4	7992	7	US-11-111-686-2
12	5315.2	55.4	7992	7	US-11-111-686-5
13	5312	55.4	7989	6	US-10-509-921-14
14	5312	55.4	7992	7	US-11-111-686-4
15	5302.2	55.3	7995	7	US-11-111-686-3
16	5127.4	53.4	7980	6	US-10-509-921-4
17	5125.8	53.4	7980	6	US-10-509-921-5
18	4987.2	52.0	7989	6	US-10-509-921-6
19	4662.4	48.6	7989	6	US-10-509-921-2
20	4327.8	45.1	7992	6	US-10-509-921-3
21	4138.4	43.1	7983	6	US-10-509-921-7
22	1558.8	16.2	2428	6	US-10-985-205-2
23	1556	16.2	2442	6	US-10-985-205-4

24	1448.2	15.1	2253	6	US-10-985-205-5	Sequence 5, Appli
25	1116.4	11.6	1935	6	US-10-985-205-7	Sequence 7, Appli
26	1116.4	11.6	1935	6	US-10-985-205-9	Sequence 9, Appli
27	1104.8	11.4	1863	6	US-10-985-205-6	Sequence 6, Appli
28	1097.6	11.4	1914	7	US-11-154-324-3	Sequence 3, Appli
29	1008.6	10.5	1746	6	US-10-985-205-8	Sequence 8, Appli
30	782.4	8.2	1338	7	US-11-117-667-1	Sequence 1, Appli
31	187.2	2.0	225	6	US-10-509-921-1	Sequence 1, Appli
32	181.8	1.9	444	6	US-10-985-205-10	Sequence 10, Appli
33	66.8	0.7	1234	7	US-11-102-240-63	Sequence 63, Appli
34	66.4	0.7	2773	7	US-11-102-240-33	Sequence 33, Appli
35	66	0.7	2773	6	US-10-689-742-41	Sequence 41, Appli
36	66	0.7	1633	6	US-10-131-826A-217	Sequence 217, App
37	66	0.7	1633	6	US-10-131-826A-453	Sequence 453, App
38	65.8	0.7	2036	6	US-10-996-217A-8	Sequence 8, Appli
39	65.6	0.7	496	6	US-10-131-826A-533	Sequence 533, App
40	65	0.7	2476	6	US-10-131-826A-489	Sequence 489, App
41	65	0.7	2487	6	US-10-689-742-165	Sequence 165, App
42	64.4	0.7	1904	6	US-10-131-826A-99	Sequence 99, Appli
43	64.4	0.7	2392	7	US-11-045-802-4	Sequence 4, Appli
44	64.4	0.7	3501	6	US-10-131-826A-37	Sequence 37, Appli
45	64	0.7	1395	6	US-10-955-054A-111	Sequence 111, App

ALIGNMENTS

RESULT 1
US-10-985-205-1
; Sequence 1, Application US/10985205
; Publication No. US20050266400A1
; GENERAL INFORMATION:
; APPLICANT: Dumonceaux, Julie
; APPLICANT: Cormier, Emmanuel G.
; APPLICANT: Gardner, Jason P.
; APPLICANT: Dragic, Tatjana
; TITLE OF INVENTION: NOVEL SEQUENCES ENCODING HEPATITIS C VIRUS GLYCOPROTEINS
; FILE REFERENCE: 71242-A/JPM/AJD
; CURRENT APPLICATION NUMBER: US/10/985,205
; CURRENT FILING DATE: 2004-11-09
; PRIOR APPLICATION NUMBER: US 60/519,536
; PRIOR FILING DATE: 2003-11-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 9599
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-10-985-205-1

Query Match	66.0%	Score	6335.8	DB	6	Length	9599
Best Local Similarity	78.9%	Pred. No.	0	Mismatches	2022	Indels	4
Matches	7573	Conservative	0	Gaps	2		
Qy	1	GCAGCCCCCTGATGGGGGCGACACTCCACATGAATCACTCCCTCTGAGGAAGTACTG	60				
Db	1	GCAGCCCCCTGATGGGGGCGACACTCCACATGAATCACTCCCTCTGAGGAAGTACTG	60				
Qy	61	TCCTCAGCAGAAAGCGCTAGCCATCGGTAGTATGATGTCGTGAGCTCCAGCTCCAGGAC	120				
Db	61	TCCTCAGCAGAAAGCGCTAGCCATCGGTAGTATGATGTCGTGAGCTCCAGGAC	120				
Qy	121	CCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG	180				
Db	121	CCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG	180				
Qy	181	GACGACCGGTCCTTTCTTGATCAACCCGCTCAATGCTGAGATTGGCGTCCCCC	240				
Db	181	GACGACCGGTCCTTTCTTGATCAACCCGCTCAATGCTGAGATTGGCGTCCCCC	240				
Qy	241	GCGAGACTGTCAGCCGAGTAGTGTGGTTCGGAAGCCCTTGTGGTACTCCCTGATAGG	300				
Db	241	GCAAGACTGTCAGCCGAGTAGTGTGGTTCGGAAGCCCTTGTGGTACTCCCTGATAGG	300				

[illegible]

3541	AAACCTTCCTGGCAACGTCGATCAATATGGGGTATGTCTGAGATCTGTCTACACGGGGCGGNA	3600
Db		
3601	CGAAGACCCCTAGCCGGTCCAAAGGTCCAAATCACCCAAATGTACACCAATGTAGACCTGG	3660
Qy		
Db		
3601	CGAGGACCATCGCATCACCAAGGGTCCGTGTCATCCAGATGTATACCAATGTGGACCAAG	3660
Qy		
3661	ACCTCGTCGGTGGCAGGCGCCCCCGGGCGGCTCCATGACACCATCAGCTAGCTGGCA	3720
Qy		
3661	ACCTTGTGGCTGGCCCGCTCCTCAAGGTTCCCGTCAATTGACACCCTGTATCCTGCGGCT	3720
Db		
3721	GCTCGGACCTTTACTTTGGTCAGGAGACATGCTGATGTCAATCCGGTGCGCCGCGAGGGC	3780
Qy		
3721	CCTCGGACCTTTACTCTGGTCAGGAGCACGCCGATGTCAATCCCGTGCGCCGCGAGGTG	3780
Db		
3781	ACAGCAGGGGAAGTCTACTCTCCCCAGGCGCGTCTCTTACTCTGAAAGGCTCCTCGGTG	3840
Qy		
3781	ATAGCAGGGGTAGCTTGTCTTCGCCCGGCCCAATTTCTTACTTTGAAAGGCTCCTCGGGG	3840
Db		
3841	GTCCATTTGTTTGCCTTTCGGGGCAGTGTGTGGCGCTCTTCGGGGTGTGTGTGCACCC	3900
Qy		
3841	GTCCGTGTTTGTCCCCCGGGGACACGCCGTGGGCGCTATTTCAGGGCGCGGTGTGCACCC	3900
Db		
3901	GGGGGTTCGGAAGGGGTGGACTTCTATACCCGTTTGAGTCTATGGAATCTTACCATCGGT	3960
Qy		
3901	GTGAGTGGCTTAAGCGGTGGACTTTTATCCCTGTGGAACCTTAGGACAACCATGAGAT	3960
Db		
3961	CTCCGGTCTTCAAGACAACCTCAACCCCCCGGCTGTATCGGAGACATTTCCAAGTGGCAC	4020
Qy		
3961	CCCCGGTGTTCAGGACAACCTCTCTCCACGACGAGTGCCTCCAGAGCTTTCAGGTGGCCC	4020
Db		
4021	ATCTGCACGCTCTACTTGGCAGCGGGAAGAGCACCAAGTGCCTGGCTGCGTATGCAAGCCC	4080
Qy		
4021	ACCTGCAATGCTCCACCGGAGCGGTAAAGACACCAAGGTGCCGGCTGCGTACGCAAGCCC	4080
Db		
4081	AAGGTTACAAGGTGCTCGTCTCTGAAACCCGTCCTGTCGGCCACTTATAGGTTTGGGGCGT	4140
Qy		
4081	AGGGCTACAAGGTGTTGGTGCTCAACCCCTCTGTGTCTGCAACGCTGGGCTTTGGTGCTT	4140
Db		
4141	ATATGTCCAAAGGCACACGGTATCGACCTTAACATCAGAACTGGGGTAAGACCAATTACCA	4200
Qy		
4141	ACATGTCCAAAGGCCATGGGGTTGATCCTTAATATCAGGACCGGGTGAGAACAAATTACCA	4200
Db		
4201	CGGGCGGCTCCATTAGTACTCCACCTATCGCAAGTTCCTTCCGACGGTGGCTGTTCTG	4260
Qy		
4201	CTGGACGCCCATCACTATCTCACTTACCGCAAGTTCCTTTCGCGAGCGGGTGCTCAG	4260
Db		
4261	GGGGCGCTATGACATCAATATGTGATGAGTGGCCACTCAACTGACTCGACTACCATCT	4320
Qy		
4261	GAGGTGCTTATGACATTAATTAATTTGTGACAGGTGCCACTCCAGGATCCACATCCATCT	4320
Db		
4321	TGGGCATCGGCACATCTCTGGACCAAGCGAGACGGTGTGAGCGCGGCTCGTGTGTCTG	4380
Qy		
4321	TGGGCATCGGCACATCTCTTGACCAAGCAGAGACTGCGGGGGGAGACTGTTGTGTCTG	4380
Db		
4381	CCACCGCTACACTCCGGGATCGGTTACGGTGCACACCCCAATATCGAGGAATAGGCC	4440
Qy		
4381	CCACTGTCTACCCCTTCGGGGTCCGTCTATGTGTCCCATCTTAACATCGAGGAGTTCCTC	4440
Db		
4441	TGTCACAATGAGAGATCCCTTCTATGGCAAGCCATCCCCCATTTGAGGCCATCAAGG	4500
Qy		
4441	TGTCCACCAACCGAGAGATCCCCCTTTTACGGAAGGCTATCCCCCTCGAGGTGATCAAG	4500
Db		
4501	GGGGGAGGCATCTCATTTTCTGCCATTTCCAAGAGAAATGTACGAGCTCGCCGCAAGC	4560
Qy		
4501	GGGGAAGACATCTCATCTCTGCGCACTCAAGAGAAAGTGCAGAGCTGCGCGGAGC	4560
Db		
4561	TGACAGGGCTCGGACTGAACGCTGTAGCATATTACGGGGGCTTGATGTCTCGTCAATAC	4620
Qy		
4561	TGTCGCAATTGGGATCAATGCGCTGGCTACTACCGCGGCTTTGACGTGTCTGTCTATCC	4620
Db		
4621	CGCCTATCGGAGACGTGTTGTTGTTGCTGGCAACAGACGCTCTAATGACGGGTTTACCGGCG	4680
Qy		
4621	CGACGAGCGGCGATGTTGTGCTGCTGTGACCCGATGTCTCATGACTGGCTTTACCGGCG	4680
Db		

4681 ATTTGACTCAGTGATCGACTGCAATATACATGTGTACCCAGACAGTTCGACTTCAGCTTGG 4740
4681 ACTTCGACTCTGTGATAGACTGCAACACGTGTGTACACTCAGACAGTTCGATTTTACGCTTG 4740
4741 ATCCACACTTCACCATTTAGAGACGACGACCGTGCCCAAGACGCGGTGTGCGCTCGCAAC 4800
4741 ACCCTACCTTTACCATTAGACACACACCGCTCCCGCCAGATGCTGTCTCCAGGACTCAAC 4800
4801 GCGAGGTAGAACTGTGCAAGGGTAGGAGTGGGATCTACAGGTTTGTGTGACTCCAGAGAAC 4860
4801 GCGGGGAGGAGTGGCAGGGGGAAGCAGGACATCTATAGATTTTGTGGACACCGGGGAGC 4860
4861 GGCCCTCGGGCATGTTTCGATTTCTCGTCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
4861 GCGCCCTCGGGCATGTTTCGACTCTGTCTGTGTGAGTGTATGACGCGGCTGTGCTT 4920
4921 GGTATGAGCTACGCGCCGTGAGACTCGGTTAGGTTGCGGGCTTACCTAAATACACAG 4980
4921 GGTATGAGCTACGCGCCGCGGAGACTACAGTTAGGCTAGAGCGGTACATGAACCCCGG 4980
4981 GGTTCGCGTGTGCGAGACCATTTGGAGTTCTGGGAGAGGCTTTTCAAGGCTTCACCC 5040
4981 GGCTTCGCGTGTGCGAGACCATTTGAAATTTGGGAGGGGCTCTTTACGGGCGCTCACTC 5040
5041 ACATAGATGCCACTTCCTGTCCGAGACTTAACAGCGAGGAGACACTTTCCTTACCCTGG 5100
5041 ATATAGATGCCACTTTTATATCCAGACAAAGCAGAGTGGGAGAACTTTTCTTACCCTGG 5100
5101 TGGCATATCAAGCTACAGTGTGCGCGAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAA 5160
5101 TAGCGTACCAAGCCACCGTGTGCGGTAGGGCTCAAGCCCTCCCGCATCTGTGGACACGA 5160
5161 TGTGGAAGTGTCTCATACGCTGAAACCTTACACTGCACGCGGCCAACACCCCTGTGTATA 5220
5161 TGTGGAAGTGTGTGATTCGCGCTTAAACCCACCTCCATGGGCCAACACCCCTGTGTATA 5220
5221 GGCTAGGAGCGCTCAAAATAGAGTTCATCTCACACACCCCATTAATAATACATCATGG 5280
5221 GACTGGGCGCTTTCAGAAATGAAGTCACTGACGACCGGCTGCTGTGTAGCGGAGTCC 5280
5281 CATGATGTCGCGCTGACCTGAGGTCGTCACTAGACACCTGGTGTGTGTAGCGGAGTCC 5340
5281 CATGATGTCGCGCGACCTGAGAGTGTCTCACAGACACCTGGTGTGTGTGGCGGCTCC 5340
5341 TTGCAAGCTTTGGCGCGATCTGCTGACGACAGGCGAGTGTGTGTATGTGGCGCAGGATCA 5400
5341 TGGCTGCTCTGGCGCGGATTTGCTGTCAACAGGCTGCGTGTGTATGTGGCGCAGGATCG 5400
5401 TCTTGTCCGGGAAGCGAGTGTGTTCCCGACAGGGAAGTCTCTACCGAGGATTCGATG 5460
5401 TCTTGTCCGGGAAGCGCGCAATATATACCTGTGACAGGAGGTTCTCTACCGAGGATTCGATG 5460
5461 AGATGAAGAGTGTGCTTCACTTCCCTTATACATCGACAGGGAATGACGCTCGCGGAGC 5520
5461 AGATGAAGAGTGTCTCAGACATTTACCGTATACGAGCAAGGATGATGCTCGCTGAGC 5520
5521 AATTCAAGCAAAAGGCGCTCGGGTGTGTGCAACCGGCCAACCAAGCAAGCGGAGGCTGTG 5580
5521 AGTTCAAGCAGAAGGCGCTCGGCTCTCGACAGCGCGTCCCGCCATGCGAGGTTATCA 5580
5581 CTCCTGTGTGAGTCCAAAGTGGCGAGCCCTTTGAGACCTTCTGGGGAAGCATGTGGA 5640
5581 CCCCTGTCTCCAGACCAACTGGCAAAACTCGAGGTCTTTTGGGCGAAGCATGTGGA 5640
5641 ATTTCAACAGCGAATACAGTACTAGAGGCTTATCCACTCTGCTGGAACCCCGCA 5700
5641 ATTTCAATCAGTGGATACAAATCTTGGCGGCTGTCAACGCTGTGCTGTGTAAACCCGCCA 5700
5701 TAGCATCATTTGATGCAATTTACAGCTTCTATCACTAGCCCGCTTCACCCCAAAACACC 5760
5701 TTGCTTCATTGATGGCTTTTACAGTGTCCGTACACAGCCCACTTAACCACTGCGCAAAACC 5760

5761 TCCTGTTTAAACATCTTTGGGGGATGGGTGGCTGCCCCAACTCGCTCTCCACAGCGCTCGT 5820
5761 TCCTTTCAACATATTGGGGGGTGGGTGGCTGCCCCAGCTCGCGCCCCCGGTCCGCGTA 5820
5821 CAGCTTTCGTGGGCGCGGCATCGCCGAGCGGCTGTGGCAGCATAGGCTTTGGGAAG 5880
5821 CTGCTTGTGGGTGTGGCTAGCTAGTGGCGCGCATCGGCAGCGTTGGACGTGGGAAG 5880
5881 TGCTCGTGGACATCTTTGGCGGGCTATGGGCGAGGGTAGCGGCGCACTCGTGCCTTTA 5940
5881 TCCTCGTGGACATCTTTGACAGGGTATGGCGGGCGTGGCGGAGCTCTTGTAGCATTTCA 5940
5941 AGTCTAGAGCGGAGGTGCGCTCCACCGAGGACCTGGTCAACTTACTCTCCCTCCCATCC 6000
5941 AGATCATAGGGGTAGGTCCCTCCACGAGGAGCTGGTCAATCTGCTGCGCCCGCATCC 6000
6001 TCTCTCTGTGGCTCGTGTGGGTGTGTGGCAGCAATACTGCTGTGGCAGCTGG 6060
6001 TCTCGCTGAGCGCTTGTAGTGGTGTGTGTCGCGAGCAATACTGCGCGCGCAGCTTG 6060
6061 GCGCGGAGAGGGGCTGTGAGTGAACCGGCTGATAGCGTTCGCTTCGCGGGGTA 6120
6061 GCGCGGCGAGGGGCGAGTGAATGAGTGAACCGGCTTAATAGCCTTCGCTCCCGGGGA 6120
6121 ACCACGCTCTCCCTTACGCACCTATGTGCTGAGAGCGAGCTGCGACGCTGTCACTCAGA 6180
6121 ACCATGTTTCCCCCAGCAGCTACGTGCGCGAGAGCGATGCGCGCGCTCACTGCCA 6180
6181 TCCTCTCTAGCTTACCATCACTCAACTGTGTGAAGCGGCTCCACAGTGGATTAATGAGG 6240
6181 TACTCAGCAGCGCTCACTGTAAACCGAGCTCTGAGCGCACTGCATCAGTGGATAAGCTCG 6240
6241 ACTGCTCTACGCCATGCTCCGGCTGTGCTTAAGGGATGTTTGGATTTGGATATGACCG 6300
6241 AGTGTACCACTCCCATGTCTCGGTTCTTGGCTAAGGGACATCTGGGACTGGATATGCGAG 6300
6301 TGTGTAGCTACTTCAAGACCTGGCTCCAGTCCAACTCTCTCGCGGTTTACCGGAGTCC 6360
6301 TGCTGAGGACTTTTAAACCTTGGCTGAAAGCAAGCTCATGCCACAACCTGCTGGATTC 6360
6361 CTTTCTGTCTATGCAACGCGGGTACAAGGGAGTCTGGCGGGGGAACGCGATCATGCAAA 6420
6361 CTTTGTGTCTCTGCGAGCGGGTATAGGGGGTCTGGCGAGGAGACGCGCATTTATGCACA 6420
6421 CCACCTGCCATGCGGAGCAGATCGCCGACATGTCAAAAACGTTCCATCGAGGATCG 6480
6421 CTCGTGCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGCGAGTGGAGTTCG 6480
6481 TAGGCGCTTAGAACCTGACAGCAACAGTGGCAACGTTCCCATCAACGATACACCA 6540
6481 TCGGTCTTAGGACTCTGAGGAACATGTGGAGTGGGACGTTCCCATTAACGCTACACCA 6540
6541 CCGGACCTTGCACACCTTCCCGCGCCCAACTATTCAGGGCGCTATGGCGGGTGGGTG 6600
6541 CCGGCGCTGTACTCTCCCTCTCTGCGCGCAACTATAAGTTCGCGCTGTGGAGGGTGTCTG 6600
6601 CTGAGGACTAGCTGGAGTGTACGCTGTGGGGATTTCCACTACGTGACGCGCATGACCA 6660
6601 CAGAGGAATACGTGGAGATAAGGGGGTGGGGACTTCCACTACGTATCGGGGTATGACTA 6660
6661 CTGACAAAGTAAAGTGCACATGCGAGGTTCCGGGCGCCCGAAATTTCTTACCGAGGTGGATG 6720
6661 CTGACAAATTTAAATGCGCGTGCAGATCCCATCGCCCGAAATTTTTCAGAAATTTGGAG 6720
6721 GAGTGGGTTCACAGGTACGCTCCGGGTGCAAACTTCTTCTACGGGAGGACGTCACT 6780
6721 GGGTGGCGCTACACAGGTTTGGCGCCCTTGGCAAGCCCTTGTGCGGGAGGAGGTATCAT 6780
6781 TCCAGGTGGGCTCAACAATACTTGGTTCGGGTTCAGCTCCCATCGGAGCCCGAACCG 6840
6781 TCAGAGTAGGACTCCACAGTACCCGGTGGGTGCGCAATTTACCTTGGAGCCCGAACCG 6840
6841 ACGTAAACAGTGTCTTACTTCCATGCTCACCGATCCCTCCCACTTACAGCAGAGACGGCTA 6900


```
Qy 9058 CATGCCTCAGAAACTTTGGGTACCAACCCTTGCGAACCTGGAGACATCGGGCCAGAAAGTG 9117
Db |||||
Qy 9061 CATGCCCTCAGAAACTTTGGGTCCGCCCTTGCAGCTTGGAGACACGGGGCCGGAGCG 9120
Db |||||
Qy 9118 TCCGCGCTAAGCTACTCTCCAGGGGGGAGGGCCGCACCTTGTGGCAGATACCTCTTTA 9177
Db |||||
Qy 9121 TCCGCGCTAGGCTTCTGTCCAGAGAGGCGAGGGCTGCCATATGTGGCAAGTACCTCTTCA 9180
Db |||||
Qy 9178 ACTGGCAGTAAGACCAAGCTTAACTCACTCAATCCCGGCCGCTGCCAGCTGGAAT 9237
Db |||||
Qy 9181 ACTGGCAGTAAGAACAAAGCTCAAACTCACTCCAAATAGCGGCCGCTGGCGGGCTGGACT 9240
Db |||||
Qy 9238 TGTCTGGCTGTTCTCGTTCGTTTACAGCGGGGAGACATATATCAAGAGCTGTCCTGCTG 9297
Db |||||
Qy 9241 TGTCCGGTGTGTTACGCGCTGGCTACAGCGGGGAGACATTTATCAAGAGTGTCCTCATG 9300
Db |||||
Qy 9298 CCCGACCCCGCTGTTTCCGTTGTGCTTACTCTTACTTCTTCTGTAGGGGTAGGCATTTACC 9357
Db |||||
Qy 9301 CCCGGCCCCGCTGTTCTGGTTCCTGCTTCTGCTCGCTGCAGGGGTAGGCATCTACC 9360
Db |||||
Qy 9358 TGTCTCCCAACCGATGAACGGGAGCTTAAACCACTCCAGGC-CTTAAGCCATTTCTGTTT 9416
Db |||||
Qy 9361 TCCTCCCAACCGATGAAGTTGGGGTAAACACTCCGGCCTCTTTAAGCCATTTCTGTTT 9420
Db |||||
Qy 9417 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 9476
Db |||||
Qy 9421 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 9480
Db |||||
Qy 9477 TTTCTTTTCTTTCTTTTAAATGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAA 9536
Db |||||
Qy 9481 TTTCTTTTCTTTCTTTTAAATGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAA 9540
Db |||||
Qy 9537 AGTCCGCTGAGCGCATGACTGCAGAGAGTCTGATGACTGSCCTCTCTGAGATCATGT 9595
Db |||||
Qy 9541 AGTCCGCTGAGCGCATGACTGCAGAGAGTCTGATAGTGGCCTCTCTGAGATCATGT 9599
Db |||||

RESULT 2
US-10-509-921-11
; Sequence 11, Application US/10509921
; Publication No. US20050250093A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Hepatitis C Virus Sub-Genomic Replicons
; FILE REFERENCE: P51335
; CURRENT APPLICATION NUMBER: US/10/509,921
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: 60/369,685
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 7979
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thepolynucleotide sequence encodes sequences from
; OTHER INFORMATION: HCV J4 Replicons
US-10-509-921-11

Query Match 64.0%; Score 6138.4; DB 6; Length 7979;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 6154; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 3416 CTTTCGCCCATCAGCGCTACTCCCAACAAACGGGGGGCTACTTGGTTGCATCATCAC 3475
Db |||||
Qy 1800 CATGGCGCTATTAGCGCTACTCCCAACAGACGGCGGCTACTTGGCTGCATCATCAC 1859
Db |||||
Qy 3476 TAGCCTCAAGCCCGGACAAAGAACAGTCCAGTCCAGGGGAGGTTCAAGTGGTTTCTACCGC 3535
Db |||||
Qy 1860 TAGCCTCAAGCCCGGACAGGAACAGGTCTGAGGGGAGGTCCAAAGTGGTCTCCACCGC 1919
Db |||||
```

```
Qy 3536 AACACAATCTTTCTGGCGACCTGSCATCAACGGCGTGTGTGTGACTGTCTTACCATGCGC 3595
Db |||||
Qy 1920 AACACAATCTTTCTGGCGACCTGCGTCAATGCGTGTGTGTGGACTGTCTTATCATGTGTC 1979
Db |||||
Qy 3596 TGGCTCGAAGACCTTAGCGGTCAAAAGGTCAATCACCCCAATGTACACCAATGTAGA 3655
Db |||||
Qy 1980 CGGCTCAAGACACCTTTCGGCGCCAAAGGGCCCAATCACCCCAATGTACACCAATGTAGA 2039
Db |||||
Qy 3656 CTTGGACCTCGTGGCTGGCAGGCGCCCCCGGGGCGCTCCATGACACCATGACAGCTG 3715
Db |||||
Qy 2040 CTTGGACCTCGTGGCTGGCAGGCGCCCCCGGGGCGCTCCATGACACCATGACAGCTG 2099
Db |||||
Qy 3716 TGGCAGCTCGGACCTTTACTTTGGTTCAGAGACATGCTGATGTCTATTCGGTGCGCCGG 3775
Db |||||
Qy 2100 TGGCAGCTCGGACCTTTACTTTGGTTCAGAGACATGCTGATGTCTATTCGGTGCGCCGG 2159
Db |||||
Qy 3776 AGGCGACAGAGGGGAAGTCTACTCTCCCGCAGGCGCTCTCTACCTGAAGGCTGCTC 3835
Db |||||
Qy 2160 AGGCGACAGAGGGGAAGTCTACTCTCCCGCAGGCGCTCTCTACCTGAAGGCTGCTC 2219
Db |||||
Qy 3836 GGGTGTTCATTTGTTTCCCTTTCGGGCGCTGCTGGGCGTCTTCCGGGCTGCTGTGTG 3895
Db |||||
Qy 2220 GGGTGTTCATTTGTTTCCCTTTCGGGCGCTGCTGGGCGTCTTCCGGGCTGCTGTGTG 2279
Db |||||
Qy 3896 CACCCGGGGGGTCCGAAAGCGGTGGAATTCATACCCGTTTGAAGTCTATGAAATACCAT 3955
Db |||||
Qy 2280 CACCCGGGGGGTCCGAAAGCGGTGGAATTCATACCCGTTTGAAGTCTATGAAATACCAT 2339
Db |||||
Qy 3956 GCGGTCTCCGGTCTTACAGACAACTCAACCCCGCCGGCTGTACCGGAGACATTTCCAAGT 4015
Db |||||
Qy 2340 GCGGTCTCCGGTCTTACAGACAACTCAACCCCGCCGGCTGTACCGGAGACATTTCCAAGT 2399
Db |||||
Qy 4016 GGCACATCTCGACGCTCTACTGCGAGCGCAAGAGCACAAAGTGCCTGCGTATGC 4075
Db |||||
Qy 2400 GGCACATCTCGACGCTCTACTGCGAGCGCAAGAGCACAAAGTGCCTGCGTATGC 2459
Db |||||
Qy 4076 AGCCCAAGGGTCAAAAGGTGCTCGTCTCGTGAACCCGTTGCCGCCACCTTTAGGGTTTG 4135
Db |||||
Qy 2460 AGCCCAAGGGTCAAAAGGTGCTCGTCTCGTGAACCCGTTGCCGCCACCTTTAGGGTTTG 2519
Db |||||
Qy 4136 GGGCTATATGTCAAAGGCAACAGGTATCGACCTTAAACATCAGAACTGGGGTGAAGACCAT 4195
Db |||||
Qy 2520 GGGCTATATGTCAAAGGCAACAGGTATCGACCTTAAACATCAGAACTGGGGTGAAGACCAT 2579
Db |||||
Qy 4196 TACCAACGGCGGCTCCATTAAGTACTCCACTGAGTGAAGTTCCTTCCGACGCTGGCTG 4255
Db |||||
Qy 2580 TACCAACGGCGGCTCCATTAAGTACTCCACTGAGTGAAGTTCCTTCCGACGCTGGCTG 2639
Db |||||
Qy 4256 TTTCTGGGGCGCTTATGACATCATATATGTGAGTGGCCACTCAACTGACTCGACTTAC 4315
Db |||||
Qy 2640 TTTCTGGGGCGCTTATGACATCATATATGTGAGTGGCCACTCAACTGACTCGACTTAC 2699
Db |||||
Qy 4316 CATCTTGGGATCGGCAAGTCTTGAACCAAGCGGAGACGGCTGGAGCGCGGCTGCTGCT 4375
Db |||||
Qy 2700 CATCTTGGGATCGGCAAGTCTTGGACCAAGCGGAGACGGCTGGAGCGCGGCTGCTGCT 2759
Db |||||
Qy 4376 GCTCGCCACCGCTACCTCGGGATCGGTTTACCGTGCACACCCCAATATCGAGGAAT 4435
Db |||||
Qy 2760 GCTCGCCACCGCTACCTCGGGATCGGTTTACCGTGCACACCCCAATATCGAGGAAT 2819
Db |||||
Qy 4436 AGGCTGTCCAAACAATGGAGAGATCCCTTCTATGGCAAAAGCCATCCCATTTAGGGCCAT 4495
Db |||||
Qy 2820 AGGCTGTCCAAACAATGGAGAGATCCCTTCTATGGCAAAAGCCATCCCATTTAGGGCCAT 2879
Db |||||
Qy 4496 CAAGGGGGGAGGAGCATCTATTTTCTGCCATTTCCAAAGAAATGTGACGAGTGCGCCG 4555
Db |||||
Qy 2880 CAAGGGGGGAGGAGCATCTATTTTCTGCCATTTCCAAAGAAATGTGACGAGTGCGCCG 2939
Db |||||
Qy 4556 AAAGCTGACAGGCTCCGAGCTGACGCTGAGCATATTTACCGGGCCTTGTATGTCTCGT 4615
Db |||||
Qy 2940 AAAGCTGACAGGCTCCGAGCTGAAACGCTGAGCATATTTACCGGGCCTTGTATGTCTCGT 2999
Db |||||
Qy 4616 CATACCGCTATCGGAGACGTCGTTGTGTCGTCGCAACAGACGCTCTTAATGACGGGTTTTCAC 4675
Db |||||
```

[illegible]

Db	4080	CGGATAGACATCATGATGGCATTTACAGTCTCTATCATAGCCCGCTCACACCAAAA	4135
Qy	5756	CACCTCTCTGTTTAAACATCTTGGGGGGATGGGTGGCTGCCAACTCGCTCTCTCCAGCGC	5815
Db	4140	CACCTCTCTGTTTAAACATCTTGGGGGGATGGGTGGCTGCCAACTCGCTCTCTCCAGCGC	4199
Qy	5816	TGCGTCAGCTTTCTGTTGGGCGCGGCATCGCCGAGCGGCTGTGGCAGATAGGCTTGG	5875
Db	4200	TGCGTCAGCTTTCTGTTGGGCGCGGCATCGCCGAGCGGCTGTGGCAGATAGGCTTGG	4259
Qy	5876	GAAGGTGCTCGTGGACATCTTGGCGGGCTATGGGGCAGGGGTAGCCGGCGCACTCGTGCG	5935
Db	4260	GAAGGTGCTCGTGGACATCTTGGCGGGCTATGGGGCAGGGGTAGCCGGCGCACTCGTGCG	4319
Qy	5936	CTTTAAAGTCAATGAGCGGCGAGGTGCCCTCCACCGAGGACCTTGTTCACTTACTCCCTGC	5995
Db	4320	CTTTAAAGTCAATGAGCGGCGAGGTGCCCTCCACCGAGGACCTTGTTCACTTACTCCCTGC	4379
Qy	5996	CATCCTCTCTCTGTTGGTGCCTTGGTCTCGGGGTCTGTGCGGAGCAATACTCGCTCGCA	6055
Db	4380	CATCCTCTCTCTGTTGGTGCCTTGGTCTCGGGGTCTGTGCGGAGCAATACTCGCTCGCA	4439
Qy	6056	CGTGGGCGCGGAGAGGGGGCTGTGACGTGGATGAACCGGCTGATAGCGTTTCGTTTCGGG	6115
Db	4440	CGTGGGCGCGGAGAGGGGGCTGTGACGTGGATGAACCGGCTGATAGCGTTTCGTTTCGGG	4499
Qy	6116	GGGTAAACCACTCTCCCTCAACGACTATGTGCTGAGAGCGACGTGCGAGCAGTGTCAAC	6175
Db	4500	GGGTAAACCACTCTCCCTCAACGACTATGTGCTGAGAGCGACGTGCGAGCAGTGTCAAC	4559
Qy	6176	TCAGATCTCTCTAGCCTTACCATCACTCAACTGCTGAAGCGGCTCCACAGTGGATTAAC	6235
Db	4560	TCAGATCTCTCTAGCCTTACCATCACTCAACTGCTGAAGCGGCTCCACAGTGGATTAAC	4619
Qy	6236	TGAGGACTGCTTACGCCATGCTCCGGCTCGTGGCTTAAGGGATGTTTGGGATTTGGATATG	6295
Db	4620	TGAGGACTGCTTACGCCATGCTCCGGCTCGTGGCTTAAGGGATGTTTGGGATTTGGATATG	4679
Qy	6296	CACGGTGTGACTGACTTCAAGACCTTGGTCTCAGTCCAACTCCTCGCGGGTTACCGGG	6355
Db	4680	CACGGTGTGACTGACTTCAAGACCTTGGTCTCAGTCCAACTCCTCGCGGGTTACCGGG	4739
Qy	6356	AGTCCCTTTCTGTTCATGCCAAACGCGGGTACAGGGAGTCTGCGCGGGGAGCGGCATCAT	6415
Db	4740	AGTCCCTTTCTGTTCATGCCAAACGCGGGTACAGGGAGTCTGCGCGGGGAGCGGCATCAT	4799
Qy	6416	GCAAAACACCTGCCCATCGGGACACAGATCGCGGACATGTCAAATAACGGTTTCCATGAG	6475
Db	4800	GCAAAACACCTGCCCATCGGGACACAGATCGCGGACATGTCAAATAACGGTTTCCATGAG	4859
Qy	6476	GATCGTAGGGCTAGAACCTTGACGAAACACGTGGCACCGAAACGTTCCCATCAACGCATA	6535
Db	4860	GATCGTAGGGCTAGAACCTTGACGAAACACGTGGCACCGAAACGTTCCCATCAACGCATA	4919
Qy	6536	CACACCGGGACCTTGACACACCTCCCGGCGGCCAATATTCACAGGCGGTATGCGGGT	6595
Db	4920	CACACCGGGACCTTGACACACCTCCCGGCGGCCAATATTCACAGGCGGTATGCGGGT	4979
Qy	6596	GGCTGCTCAGGAGTACGTGGAGTTTACGCTGTGGGGATTTCCATCTACGTGACGGGCAT	6655
Db	4980	GGCTGCTCAGGAGTACGTGGAGTTTACGCTGTGGGGATTTCCATCTACGTGACGGGCAT	5039
Qy	6656	GACCACTGACAAACGTTAAAGTGCCCATGCCAGTTTCGGGCCCGGAAATTTCTTCAGGAGT	6715
Db	5040	GACCACTGACAAACGTTAAAGTGCCCATGCCAGTTTCGGGCCCGGAAATTTCTTCAGGAGT	5099
Qy	6716	GGATGGAGTGGTGTGACAGGTTACGCTTCGGCGGTGCAAAACCTTCTCTACGGGAGNCGT	6775
Db	5100	GGATGGAGTGGTGTGACAGGTTACGCTTCGGCGGTGCAAAACCTTCTCTACGGGAGNCGT	5159
Qy	6776	CACGTTCCAGTTCGGGCTCAACCAATACTTGGTTCGGGTTCGAGCTCCCATTCGAGCCCGGA	6835
Db	5160	CACGTTCCAGTTCGGGCTCAACCAATACTTGGTTCGGGTTCGAGCTCCCATTCGAGCCCGGA	5219

6836 ACCGACGTAACAGTGTCTTACTTCCATGCTCACGATCCCTCCCACTTTACAGCAGAC 6895
| | | | |
Db | | | | |
5220 ACCGACGTAACAGTGTCTTACTTCCATGCTCACGATCCCTCCCACTTTACAGCAGAC 5279
| | | | |
Qy | | | | |
6896 GGCTAAGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTTAGCGAGCTCATCAGCTAGCCA 6955
| | | | |
Db | | | | |
5280 GGCTAAGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTTAGCGAGCTCATCAGCTATCCA 5339
| | | | |
Qy | | | | |
6956 GTTGTCTGGCTTCTTTGAAGCGACATGCACTACCCACCACTGACTCCCGCGAGCTGA 7015
| | | | |
Db | | | | |
5340 GTTGTCTGGCTTCTTTGAAGCGACATGCACTACCCACCACTGACTCCCGCGAGCTGA 5399
| | | | |
Qy | | | | |
7016 CCTCATCGAGGCCAACCCTTTGTGCGCGCAGAGATGGCGGAAACATCACTCGCGTGA 7075
| | | | |
Db | | | | |
5400 CCTCATCGAGGCCAACCCTTTGTGCGCGCAGAGATGGCGGAAACATCACTCGCGTGA 5459
| | | | |
Qy | | | | |
7076 GTACAGAGTAATAGGTAGTAATTTCTGGACTCTTTTCGAACCGCTTTCACGCGGAGGGGATGA 7135
| | | | |
Db | | | | |
5460 GTACAGAGTAATAGGTAGTAATTTCTGGACTCTTTTCGAACCGCTTTCACGCGGAGGGGATGA 5519
| | | | |
Qy | | | | |
7136 GAGGAGATATCCGTGCGGCGGAGATCTCGGAAAAATCCAGGAAGTTCCCTCAGCGTT 7195
| | | | |
Db | | | | |
5520 GAGGAGATATCCGTGCGGCGGAGATCTCGGAAAAATCCAGGAAGTTCCCTCAGCGTT 5579
| | | | |
Qy | | | | |
7196 GCCCATATGGGCACGCCCGGACTCAATCTCCACTGCTAGAGTCTCGGAGGACCCGGA 7255
| | | | |
Db | | | | |
5580 GCCCATATGGGCACGCCCGGACTCAATCTCCACTGCTAGAGTCTCGGAGGACCCGGA 5639
| | | | |
Qy | | | | |
7256 CTACGTCCCTCCGCTGGTACAGGATGCCATTTGCCACTTCAAGGCTCTCCCAATACC 7315
| | | | |
Db | | | | |
5640 CTACGTCCCTCCGCTGGTACAGGATGCCATTTGCCACTTCAAGGCTCTCCCAATACC 5699
| | | | |
Qy | | | | |
7316 ACCTCCAGGAGAAAGAGACGGTGTCTCTGACAGAAATCCAAATGTGTCTCTGCTTGGC 7375
| | | | |
Db | | | | |
5700 ACCTCCAGGAGAAAGAGACGGTGTCTCTGACAGAAATCCAAATGTGTCTCTGCTTGGC 5759
| | | | |
Qy | | | | |
7376 GAGCTCGCCACTAGACCTTCGGTAGCTCCGGATCGTGGCGGTGTGATAGCGGCACGCG 7435
| | | | |
Db | | | | |
5760 GAGCTCGCCACTAGACCTTCGGTAGCTCCGGATCGTGGCGGTGTGATAGCGGCACGCG 5819
| | | | |
Qy | | | | |
7436 GACCGCCCTTCCTGACCTGGCTCCGACGACGGTGACAAAGGATCCGAGTTGAGTCTGTA 7495
| | | | |
Db | | | | |
5820 GACCGCCCTTCCTGACCTGGCTCCGACGACGGTGACAAAGGATCCGAGTTGAGTCTGTA 5879
| | | | |
Qy | | | | |
7496 CTCCTCCATGCCCCCTTTGAAGGGAGCGCGGGGACCCCGATCTCAGCGACGGGTCTTG 7555
| | | | |
Db | | | | |
5880 CTCCTCCATGCCCCCTTTGAAGGGAGCGCGGGGACCCCGATCTCAGCGACGGGTCTTG 5939
| | | | |
Qy | | | | |
7556 GTCTACCGTAGTGAGGAGCTAGTGAGGATCTGCTGCTGCTCAATGTCTATACGTTG 7615
| | | | |
Db | | | | |
5940 GTCTACCGTAGTGAGGAGCTAGTGAGGATGTGCTGCTGCTCAATGTCTATACGTTG 5999
| | | | |
Qy | | | | |
7616 GACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAAGTAAGCTGCCCATCAACCCGTT 7675
| | | | |
Db | | | | |
6000 GACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAAGTAAGCTGCCCATCAACCCGTT 6059
| | | | |
Qy | | | | |
7676 GAGCAACTCTTTGCTGCGTCAACAACATGTGTTACGCCCAACAACATCCCGCAGCGCAAG 7735
| | | | |
Db | | | | |
6060 GAGCAACTCTTTGCTGCGTCAACAACATGTGTTACGCCCAACAACATCCCGCAGCGCAAG 6119
| | | | |
Qy | | | | |
7736 CCTCCGGCAGAGAGGTCACTTTGACAGATTGCAAGTCTCGGATGATCATTCGGGA 7795
| | | | |
Db | | | | |
6120 CCTCCGGCAGAGAGGTCACTTTGACAGATTGCAAGTCTCGGATGATCATTCGGGA 6179
| | | | |
Qy | | | | |
7796 CGTACTCAAGGAGATGAAGCGAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGA 7855
| | | | |
Db | | | | |
6180 CGTACTCAAGGAGATGAAGCGAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGA 6239
| | | | |
Qy | | | | |
7856 GGAGGCTCGAAGCTGACGCCGCCCAATTCTGGGCCAAATCCAAATTTGGTCTATGGGGCAAA 7915
| | | | |
Db | | | | |
6240 GGAGGCTCGAAGCTGACGCCGCCCAATTCTGGGCCAAATCCAAATTTGGTCTATGGGGCAAA 6299
| | | | |

7916 GGAGTCCGGAACCTATCCAGCAGGGCGGTTAAACCATCCGCTCGTGTGGAGGACTT 7975
| | | | |
Db | | | | |
6300 GGAGTCCGGAACCTATCCAGCAGGGCGGTTAAACCATCCGCTCGTGTGGAGGACTT 6359
| | | | |
Qy | | | | |
7976 GCTGGAAGACACTGAAAACCAATTGACACCACTCATGSCAAAAAGTGTGTTCTG 8035
| | | | |
Db | | | | |
6360 GCTGGAAGACACTGAAAACCAATTGACACCACTCATGSCAAAAAGTGTGTTCTG 6419
| | | | |
Qy | | | | |
8036 CGTCCAAACAGAGAGGAGGCGCAAGCAGCTCGCTTATCTATTCCAGACCTCGG 8095
| | | | |
Db | | | | |
6420 CGTCCAAACAGAGAGGAGGCGCAAGCAGCTCGCTTATCTATTCCAGACCTCGG 6479
| | | | |
Qy | | | | |
8096 AGTTCTGTGTATGCGAGAGATGGCCCTTTTACGAGCTGTCTCCACCTCTCTCAGGCCGT 8155
| | | | |
Db | | | | |
6480 AGTTCTGTGTATGCGAGAGATGGCCCTTTTACGAGCTGTCTCCACCTCTCTCAGGCCGT 6539
| | | | |
Qy | | | | |
8156 GATGGGTCTCTCATACGGATTTCAATCTCCCCCAAGCAGCGGTCTGAGTTCTCTGTGAA 8215
| | | | |
Db | | | | |
6540 GATGGGTCTCTCATACGGATTTCAATCTCCCCCAAGCAGCGGTCTGAGTTCTCTGTGAA 6599
| | | | |
Qy | | | | |
8216 TACCTGGAAATCAAGAAATCGCCTATGCGGTCTCTCATGACACCCGCTGTTTGTACTC 8275
| | | | |
Db | | | | |
6600 TACCTGGAAATCAAGAAATCGCCTATGCGGTCTCTCATGACACCCGCTGTTTGTACTC 6659
| | | | |
Qy | | | | |
8276 AACGTCACTGAGAGTGACATTCGTGTGTGAGAGTCAATTTACCAATGTGTGTACTTGGC 8335
| | | | |
Db | | | | |
6660 AACGTCACTGAGAGTGACATTCGTGTGTGAGAGTCAATTTACCAATGTGTGTACTTGGC 6719
| | | | |
Qy | | | | |
8336 CCCGAGGCCAGACAGGCCATAAGGTCTCTCAGAGCGGCTTTACATCGGGGGTCCCTT 8395
| | | | |
Db | | | | |
6720 CCCGAGGCCAGACAGGCCATAAGGTCTCTCAGAGCGGCTTTACATCGGGGGTCCCTT 6779
| | | | |
Qy | | | | |
8396 GACTAACTCAAAAGGGCAGAACTCGGTTATCGCGGTGCGCGCAAGTGGCGTCTGCAC 8455
| | | | |
Db | | | | |
6780 GACTAACTCAAAAGGGCAGAACTCGGTTATCGCGGTGCGCGCAAGTGGCGTCTGCAC 6839
| | | | |
Qy | | | | |
8456 GACTAGCTGCGGTAAATACCTCACAATGTACTTGAAGGCCACTGACAGCTGTGCGAGTGC 8515
| | | | |
Db | | | | |
6840 GACTAGCTGCGGTAAATACCTCACAATGTACTTGAAGGCCACTGACAGCTGTGCGAGTGC 6899
| | | | |
Qy | | | | |
8516 AAAGTCCAGACTGACAGTGTCTGTAACGGAGACGACTTGTGCTGTTATCTGTGAAG 8575
| | | | |
Db | | | | |
6900 AAAGTCCAGACTGACAGTGTCTGTAACGGAGACGACTTGTGCTGTTATCTGTGAAG 6959
| | | | |
Qy | | | | |
8576 CGCGGAACCCAGAGGATCGGGCGCTTACGAGCTTACGAGGCTATGACTAGGTA 8635
| | | | |
Db | | | | |
6960 CGCGGAACCCAGAGGATCGGGCGCTTACGAGCTTACGAGGCTATGACTAGGTA 7019
| | | | |
Qy | | | | |
8636 TTCCGCCCCCCCCCGGGGATCCGCCCAACAGAAATACGACCTGGAGCTGATAACATCATG 8695
| | | | |
Db | | | | |
7020 TTCCGCCCCCCCCCGGGGATCCGCCCAACAGAAATACGACCTGGAGCTGATAACATCATG 7079
| | | | |
Qy | | | | |
8696 TTCTCTCAATGTGTGTCAGTCTGGCGACGATGCAATCTGGCAAAAGGGTATACTACTCACCCG 8755
| | | | |
Db | | | | |
7080 TTCTCTCAATGTGTGTCAGTCTGGCGACGATGCAATCTGGCAAAAGGGTATACTACTCACCCG 7139
| | | | |
Qy | | | | |
8756 TGACCCCAACCAACCCCTTGCACGGCTCGGTGGGAGACAGCTAGACACCTCCCAATCAA 8815
| | | | |
Db | | | | |
7140 TGACCCCAACCAACCCCTTGCACGGCTCGGTGGGAGACAGCTAGACACCTCCCAATCAA 7199
| | | | |
Qy | | | | |
8816 CTCTTGGCTAGGCAATATCATGTATGCGCCCAACCTTATGGCAAGGATGATTTCTGAT 8875
| | | | |
Db | | | | |
7200 CTCTTGGCTAGGCAATATCATGTATGCGCCCAACCTTATGGCAAGGATGATTTCTGAT 7259
| | | | |
Qy | | | | |
8876 GACTCACTTTTCTCCATCTTTAGTCTCAAGACAACTTGAAGAAAGCCCTCGATTGTCA 8935
| | | | |
Db | | | | |
7260 GACTCACTTTTCTCCATCTTTAGTCTCAAGACAACTTGAAGAAAGCCCTCGATTGTCA 7319
| | | | |
Qy | | | | |
8936 GACTCACTCGGGCTTCTACTCCATTTGAGCCACTTGCCTACCTCAGATCATTTGAACGACT 8995
| | | | |
Db | | | | |
7320 GACTCACTCGGGCTTCTACTCCATTTGAGCCACTTGCCTACCTCAGATCATTTGAACGACT 7379
| | | | |
Qy | | | | |
8996 CCATGGTCTTAGCGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGC 9055
| | | | |

Db	7380	CCATGGTCTTAGCGCATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGC	7439	
Qy	9056	TTTCATGCCTCAGAAACTTGGGGTACCACCTTTGGGAACCTCGAGACATCGGGCGCAGAG	9115	
Db	7440	TTTCATGCCTCAGAAACTTGGGGTACCACCTTTGGGAACCTCGAGACATCGGGCGCAGAG	7499	
Qy	9116	TGTCCGCGCTAAGCTACTCTCCAGGGGGGAGGGCGGCCACTTTGTGGCAGATACCTCTT	9175	
Db	7500	TGTCCGCGCTAAGCTACTCTCCAGGGGGGAGGGCGGCCACTTTGTGGCAGATACCTCTT	7559	
Qy	9176	TAACTGGGCAGTAAGAACCAAGTTAACTCACTCAATCCCGCGCGCTCCAGCTGGA	9235	
Db	7560	TAACTGGGCAGTAAGAACCAAGTTAACTCACTCAATCCCGCGCGCTCCAGCTGGA	7619	
Qy	9236	CTTGCTGCGCTGTTCTCGTCTGCTGTGTACAGCGGGGAGACATATATACAGAGCTGTCG	9295	
Db	7620	CTTGCTGCGCTGTTCTCGTCTGCTGTGTACAGCGGGGAGACATATATACAGAGCTGTCG	7679	
Qy	9296	TGCCCCGACCCCGCTGTTCCGTTGTGCTTACTCTTCTTCTGTAGGGGTAGGCATTTA	9355	
Db	7680	TGCCCCGACCCCGCTGTTCCGTTGTGCTTACTCTTCTTCTGTAGGGGTAGGCATTTA	7739	
Qy	9356	CTGTCTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCTTAAAGCCATTTCTCTGT	9415	
Db	7740	CTGTCTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCTTAAAGCCATTTCTCTGT	7799	
Qy	9416	TT	9475	
Db	7800	TT	7859	
Qy	9476	CTTCTCTTTTCCCTCTTTAATGGTGGCTCCATCTTAGCCCTAGTCAACGGCTAGCTGTGA	9535	
Db	7860	CTTCTCTTTTCCCTCTTTAATGGTGGCTCCATCTTAGCCCTAGTCAACGGCTAGCTGTGA	7919	
Qy	9536	AAGGTCGGTAGCCGATGATCCAGAGAGTGTGTATGTGAGTCTCTTGCAGATCATGT	9595	
Db	7920	AAGGTCGGTAGCCGATGATCCAGAGAGTGTGTATGTGAGTCTCTTGCAGATCAAGT	7979	

RESULT 3

US-10-509-921-12

; Sequence 12, Application US/10509921

; Publication No. US20050250093A1

; GENERAL INFORMATION:

; APPLICANT: SmithKline Beecham Corporation

; TITLE OF INVENTION: Hepatitis C Virus Sub-Genomic Replicons

; FILE REFERENCE: P51335

; CURRENT APPLICATION NUMBER: US/10/509,921

; CURRENT FILING DATE: 2004-10-01

; PRIOR APPLICATION NUMBER: 60/369,685

; PRIOR FILING DATE: 2002-04-03

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 7979

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: The polynucleotide sequence encodes sequences from HCV J4 Replicons

US-10-509-921-12

Query Match 64.0%; Score 6136.8; DB 6; Length 7979;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 6153; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy	3416	CTTTGGCGCCATACGGCTACTCCCAACAAACGGGGGGCGTACTTGGTTGCATCATCAC	3475	
Db	1800	CATGGCGCTATTACGGCTACTCCCAACAGACGCGAGGCGCTACTTGGCTGCATCATCAC	1859	
Qy	3476	TAGCTCACAGCCGGGACAGACCAAGCTCGAAGGGAGGTTCAAGTGGTTTCTACCGC	3535	

Db	1860	TAGCTCACAGCCGGGACAGAACCCAGGTCCAGGGGGAGGTCCAAGTGGTCTCCACCGC	1919	
Qy	3536	AACACAACTTTCTTGGCGAAGCTGATCAACCGGCTGTGTGGAGTGTCTACCATGGCGC	3595	
Db	1920	AACACAACTTTCTTGGCGAAGCTGATCAACCGGCTGTGTGGAGTGTCTATCATGGTGC	1979	
Qy	3596	TGGCTCGAAGACCCCTAGCCGCTCAAAAGGTCCAATCACCCAAATGTACACCAATGTAGA	3655	
Db	1980	CGGCTCAAAAGACCCCTTGGCGCCAAAGGGCCCAATCACCCAAATGTACACCAATGTAGA	2039	
Qy	3656	CTTGACCTCTGTCGGCTGGCAGCGCCCCCGGGGCGGCTCCATCAACCAATGACAGCTG	3715	
Db	2040	CTTGACCTCTGTCGGCTGGCAGCGCCCCCGGGGCGGCTCCAAGACCAATGACAGCTG	2099	
Qy	3716	TGGCAGCTCGGACCTTTACTTTGCTCAGGAGACATGCTGATGTCTCCGGTGGCGCGG	3775	
Db	2100	TGGCAGCTCGGACCTTTACTTTGCTCAGGAGACATGCTGATGTCTCCGGTGGCGCGG	2159	
Qy	3776	AGGCGACAGCAGGGGAAGTCTACTCTCCCGAGGCGGCTCTCTACCTCAAGAGCTCTC	3835	
Db	2160	AGGCGACAGCAGGGGAAGTCTACTCTCCCGAGGCGGCTCTCTACCTCAAGAGCTCTC	2219	
Qy	3836	GGTGTGCTCATTTGCTTTGCGCTTCGGGCGACGCTGCTGGGCGCTCTTCGGGCTGCTGTG	3895	
Db	2220	GGTGTGCTCATTTGCTTTGCGCTTCGGGCGACGCTGCTGGGCGCTCTTCGGGCTGCTGTG	2279	
Qy	3896	CACCGGGGGGTGCGGAAGCGGTGACTTATATACCCGTTGAGTCTATGGAACTACCAT	3955	
Db	2280	CACCGGGGGGTGCGGAAGCGGTGACTTATATACCCGTTGAGTCTATGGAACTACCAT	2339	
Qy	3956	GGGCTCTCGGCTTTCAGAGACAACTCAACCCCCCGGCTGTACCGACAGACATTCCAAGT	4015	
Db	2340	GGGCTCTCGGCTTTCAGAGACAACTCAACCCCCCGGCTGTACCGACAGACATTCCAAGT	2399	
Qy	4016	GGCAGATCTGACGCTCTTACTTGGCAGCGGCAAGAGCACAAAGTCCGGCTCGTATGC	4075	
Db	2400	GGCAGATCTGACGCTCTTACTTGGCAGCGGCAAGAGCACAAAGTCCGGCTCGTATGC	2459	
Qy	4076	AGCCCAAGGGTACAAGGTGCTCGTCTGAAACCGCTCCGTTGCGCGCACCTTAGGGTTGG	4135	
Db	2460	AGCCCAAGGGTACAAGGTGCTCGTCTGAAACCGCTCCGTTGCGCGCACCTTAGGGTTGG	2519	
Qy	4136	GGCGTATATGTCCAAGGCAACCGGTATGACCCCTAAACATCAGAACCTGGGGTAAGGACAT	4195	
Db	2520	GGCGTATATGTCCAAGGCAACCGGTATGACCCCTAAACATCAGAACCTGGGGTAAGGACAT	2579	
Qy	4196	TACCAAGGGGCTTCAATACGCTTACCTTACCTATGCGAGTCTCTTGGCGGCGGTGCTG	4255	
Db	2580	TACCAAGGGGCTTCAATACGCTTACCTTACCTATGCGAGTCTCTTGGCGGCGGTGCTG	2639	
Qy	4256	TTCTGGGGCGGCTATGACATATAATATGTATGATGAGTGCCACTCAACTGACCTGATAC	4315	
Db	2640	TTCTGGGGCGGCTATGACATATAATATGTATGATGAGTGCCACTCAACTGACCTGATAC	2699	
Qy	4316	CATCTTTGGGCATCGGCACAGTCTTGACCAAGCGGAGACGGCTGGAGCGGCTGCTCGT	4375	
Db	2700	CATCTTTGGGCATCGGCACAGTCTTGACCAAGCGGAGACGGCTGGAGCGGCTGCTCGT	2759	
Qy	4376	GCTCGGCACCGCTACACCTCCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAAT	4435	
Db	2760	GCTCGGCACCGCTACACCTCCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAAT	2819	
Qy	4436	AGGCTGTCCAACAATGGAGAGATCCCTTCTATATGGCAAGCCATCCCAATGAGGCGAT	4495	
Db	2820	AGGCTGTCCAACAATGGAGAGATCCCTTCTATATGGCAAGCCATCCCAATGAGGCGAT	2879	
Qy	4496	CAAGGGGGGAGGCGATCTCAATTTTCTGCCATTTCCAAGAAGAAATGTACGAGCTCGCGC	4555	
Db	2880	CAAGGGGGGAGGCGATCTCAATTTTCTGCCATTTCCAAGAAGAAATGTACGAGCTCGCGC	2939	
Qy	4556	AAAGCTGACAGGCGCTCGGACTGAAACGCTGTAGCATATATACGGGGCGCTTGTATGTCGT	4615	
Db	2940	AAAGCTGACAGGCGCTCGGACTGAAACGCTGTAGCATATATACGGGGCGCTTGTATGTCGT	2999	

Qy 4616 CATACCGCTATCGAGACGTCGTTGTCGTGGCAACAGACGCTCTAAATGACGGGTTTCAC 4675
Db 3000 CATACCGCTATCGGAGACGTCGTTGTCGTGGCAACAGACGCTCTAATGACGGGTTTCAC 3059
Qy 4676 CGGCAATTTGACTCAGTGATCGATGCAATACATGTGTCTCCAGACAGTCGACTTCAG 4735
Db 3060 CGGCAATTTGACTCAGTGATCGACTGCAATACATGTGTCTCCAGACAGTCGACTTCAG 3119
Qy 4736 CTTGGATCCCACTTCACATTTAGACGACGACGACGTCGCCCAAGACGCGGTGTGCGGCTC 4795
Db 3120 CTTGGATCCCACTTCACATTTAGACGACGACGACGTCGCCCAAGACGCGGTGTGCGGCTC 3179
Qy 4796 GCAACGGCGAGGTAGAACTGCGCAGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGG 4855
Db 3180 GCAACGGCGAGGTAGAACTGCGCAGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGG 3239
Qy 4856 AGAAGCGGCTCGGGCATGTTTCGATTCTTCGGTCTCGTGTGAGTGTCTATGACGGGGCTG 4915
Db 3240 AGAAGCGGCTCGGGCATGTTTCGATTCTTCGGTCTCGTGTGAGTGTCTATGACGGGGCTG 3299
Qy 4916 TCGTTGGTATGAGCTCAGCGCGCTGAGACCTCGGTTAGGTTGCGGGCTTACCTAAATAC 4975
Db 3300 TCGTTGGTATGAGCTCAGCGCGCTGAGACCTCGGTTAGGTTGCGGGCTTACCTAAATAC 3359
Qy 4976 ACCAGGTTGCCGCTGCCAGGACCATCTGAGTTCTGGGAGAGCGCTTTCACAGGCT 5035
Db 3360 ACCAGGTTGCCGCTGCCAGGACCATCTGAGTTCTGGGAGAGCGCTTTCACAGGCT 3419
Qy 5036 CACCCACATAGATGCCACTTCTCTGCCAGACTAAACAGGACGAGACAACTTTTCCTTA 5095
Db 3420 CACCCACATAGATGCCACTTCTCTGCCAGACTAAACAGGACGAGACAACTTTTCCTTA 3479
Qy 5096 CCGTGTGGCATATCAAGCTACAGTGTGCGCCAGGCTCAAGCTCCACTCCATCGTGGGA 5155
Db 3480 CCGTGTGGCATATCAAGCTACAGTGTGCGCCAGGCTCAAGCTCCACTCCATCGTGGGA 3539
Qy 5156 CCAAATGTGGAGTGTCTCATACGCTGAACTACACTGACAGGGCCCAACCCCTGCT 5215
Db 3540 CCAAATGTGGAGTGTCTCATACGCTGAACTACACTGACAGGGCCCAACCCCTGCT 3599
Qy 5216 GTATAGGCTAGAGCGCTCCAAATAGAGTCACTCTCCACACACCCCAATAATACAT 5275
Db 3600 GTATAGGCTAGAGCGCTCCAAATAGAGTCACTCTCCACACCCCAATAATACAT 3659
Qy 5276 CATGCAATGATGTGCGGTGACCTGGAGGTGTCTACATAGCACCTGGGTGTGGTAGGCGG 5335
Db 3660 CATGCAATGATGTGCGGTGACCTGGAGGTGTCTACATAGCACCTGGGTGTGGTAGGCGG 3719
Qy 5336 AGTCCTTCGAGCTTTGGCGGCATATGCTGTGACGACAGGAGTGTGTCATTTGGGCGAG 5395
Db 3720 AGTCCTTCGAGCTTTGGCGGCATATGCTGTGACGACAGGAGTGTGTCATTTGGGCGAG 3779
Qy 5396 GATCATCTTGTCCGGGAAGCGAGCTGTGCTTCCGACAGGGAAGTCTCTACAGAGATT 5455
Db 3780 GATCATCTTGTCCGGGAAGCGAGCTGTGCTTCCGACAGGGAAGTCTCTACAGAGATT 3839
Qy 5456 CGATGAGTGGAGAGTGTGCTCAAACTTCCTTACATPCGACGAGGAATGCACTCGC 5515
Db 3840 CGATGAGTGGAGAGTGTGCTCAAACTTCCTTACATPCGACGAGGAATGCACTCGC 3899
Qy 5516 CGAGCAATTCAGCAAAAGCGCTCGGTTGTGCAAGCGGCACCAAGCAAGCGGAGGC 5575
Db 3900 CGAGCAATTCAGCAAAAGCGCTCGGTTGTGCAAGCGGCACCAAGCAAGCGGAGGC 3959
Qy 5576 TCGTCTCCCGTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACTTCTGGGCGAAGCAAT 5635
Db 3960 TCGTCTCCCGTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACTTCTGGGCGAAGCAAT 4019
Qy 5636 GTGGAAATTTTCACGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAAACCC 5695
Db 4020 GTGGAAATTTTCACGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAAACCC 4079

Qy 5696 CGGATAGCATCATTTGATGCAATTTATGAGCTTCTATCAGCTTCTATCAGCTAGCCCGCTACACCCCAAAA 5755
Db 4080 CGGATAGCATCATTTGATGCAATTTATGAGCTTCTATCAGCTTCTATCAGCTAGCCCGCTACACCCCAAAA 4139
Qy 5756 CACCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAACTCGCTCTCTCCAGGCG 5815
Db 4140 CACCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAACTCGCTCTCTCCAGGCG 4199
Qy 5816 TGGTCAAGCTTTCGTGGGCGCCGSCATCGCGGAGCGGCTGTTGGCAGCATPAGGCTTGG 5875
Db 4200 TGGTCAAGCTTTCGTGGGCGCCGSCATCGCGGAGCGGCTGTTGGCAGCATPAGGCTTGG 4259
Qy 5876 GAAGGTCTCTGTGACATCTTTGGCGGCTATGGGCGCAGGGGTATGCCGCGCACTCGTGGC 5935
Db 4260 GAAGGTCTCTGTGACATCTTTGGCGGCTATGGGCGCAGGGGTATGCCGCGCACTCGTGGC 4319
Qy 5936 CTTTAAAGGTCAATGAGCGGCGAGGTGCCCTCCACCGAGGACCTGTGTAACCTTACTCCCTGC 5995
Db 4320 CTTTAAAGGTCAATGAGCGGCGAGGTGCCCTCCACCGAGGACCTGTGTAACCTTACTCCCTGC 4379
Qy 5996 CATCTCTCTCTGCTGCTCGGCTCGTGGGCTCGTGGCAGCAATACTGCGTCCGCA 6055
Db 4380 CATCTCTCTCTGCTGCTCGGCTCGTGGGCTCGTGGCAGCAATACTGCGTCCGCA 4439
Qy 6056 CGTGGGCGCGGAGAGGGGCTGTGCACTGATGATGAACCGGCTGATAGCGTTTCGCTTCGCG 6115
Db 4440 CGTGGGCGCGGAGAGGGGCTGTGCACTGATGATGAACCGGCTGATAGCGTTTCGCTTCGCG 4499
Qy 6116 GGGTAAACACGCTCTCCCTACGCACTATGCTGTGAGAGGACGCTGACGACGCTGTCTAC 6175
Db 4500 GGGTAAACACGCTCTCCCTACGCACTATGCTGTGAGAGGACGCTGACGACGCTGTCTAC 4559
Qy 6176 TCAGATCTCTCTAGCTTACCACATCACTCACTGCTGAAGCGGCTCCACAGTGATTA 6235
Db 4560 TCAGATCTCTCTAGCTTACCACATCACTCACTGCTGAAGCGGCTCCACAGTGATTA 4619
Qy 6236 TGAGGACTGCTCTACGCACTGCTCCGCTCGTGGCTAAGGATGTTTGGGATTTGGATG 6295
Db 4620 TGAGGACTGCTCTACGCACTGCTCCGCTCGTGGCTAAGGATGTTTGGGATTTGGATG 4679
Qy 6296 CACGGTGTGATCACTCAAGACTGCTGCTCAAGTCCAGTCCAACTCTCTGCGCGGTTACCGGG 6355
Db 4680 CACGGTGTGATCACTCAAGACTGCTGCTCAAGTCCAGTCCAACTCTCTGCGCGGTTACCGGG 4739
Qy 6356 AGTCCTCTCTGTCATGCCAAACGGGGTACAAGGGAGTCTGCGGGGGGACGCGCATCAT 6415
Db 4740 AGTCCTCTCTGTCATGCCAAACGGGGTACAAGGGAGTCTGCGGGGGGACGCGCATCAT 4799
Qy 6416 GCAAAACACCTGCCCATGCGGAGCACAGATCGCCGACATGTCAAAAACGGTTCCATGAG 6475
Db 4800 GCAAAACACCTGCCCATGCGGAGCACAGATCGCCGACATGTCAAAAACGGTTCCATGAG 4859
Qy 6476 GATCGTAGGGCTTAGAACCTGACAGCAACGTTGGCAAGAAAGTTCCTCCATCAACGCAATA 6535
Db 4860 GATCGTAGGGCTTAGAACCTGACAGCAACGTTGGCAAGAAAGTTCCTCCATCAACGCAATA 4919
Qy 6536 CACCAACGGACTTGTGCAACCCCTCCCGCGCCCAACTATTTCCAGGGCGCTATGCGGGT 6595
Db 4920 CACCAACGGACTTGTGCAACCCCTCCCGCGCCCAACTATTTCCAGGGCGCTATGCGGGT 4979
Qy 6596 GGCTGCTGAGGAGTACGTGGAGGTTTACGCGTGTGGGGATTTCCACTACTGACGCGGAT 6655
Db 4980 GGCTGCTGAGGAGTACGTGGAGGTTTACGCGTGTGGGGATTTCCACTACTGACGCGGAT 5039
Qy 6656 GACCACTGACAAACGTTAAAGTCCCATGCAAGTTCCGGCCCCCGAAATTTCTTCAACGAGGT 6715
Db 5040 GACCACTGACAAACGTTAAAGTCCCATGCAAGTTCCGGCCCCCGAAATTTCTTCAACGAGGT 5099
Qy 6716 GGATGAGTGGGTTGCAACGAGTACGCTCCCGCTGCAAACTCTTCTTACGGGAGGAGT 6775
Db 5100 GGATGAGTGGGTTGCAACGAGTACGCTCCCGCTGCAAACTCTTCTTACGGGAGGAGT 5159
Qy 6776 CACGTTCCAGGTCGGGCTCAACCAATACTTGGTGGGTCGAGCTCCCATGCGAGCCCGA 6835

Db	5160		CACGTTCCAGGTGGGCTCAACCAATACTTGGTCGGGTCGCAGCTCCCATGCGAGCCGA	5219	
Qy	6836		ACCGGACGTAAACAGTCTTACTTCCATGCTCACCAGTCCCTCCCAATTTACAGCAGAC	6895	
Db	5220		ACCGGACGTAAACAGTCTTACTTCCATGCTCACCAGTCCCTCCCAATTTACAGCAGAC	5279	
Qy	6896		GGCTAAGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCA	6955	
Db	5280		GGCTAAGCGTAGGCTGGCTAGAGGGTCTCCCCCTCTTTAGCCAGCTCATCAGCTATCCA	5339	
Qy	6956		GTTGTCTGGCCCTTTTGAAGCGCATGCACTACCCACCATGACTCCCGGACCGCTGA	7015	
Db	5340		GTTGTCTGGCCCTTTTGAAGCGCATGCACTACCCACCATGACTCCCGGACCGCTGA	5399	
Qy	7016		CCTCATGAGGCCAACTCTTGGCGGCGAGGAGATGGCGGAAACATCACTCGCGTGA	7075	
Db	5400		CCTCATGAGGCCAACTCTTGGCGGCGAGGAGATGGCGGAAACATCACTCGCGTGA	5459	
Qy	7076		GTGAGAGAAATAGGTAGTAAATCTGGACTCTTTGAAACCGCTTTCAGCGGAGGGGATGA	7135	
Db	5460		GTGAGAGAAATAGGTAGTAAATCTGGACTCTTTGAAACCGCTTTCAGCGGAGGGGATGA	5519	
Qy	7136		GAGGAGATATCGTTCGCGCGGAGATCTTCGAAATCCAGGAAGTTCCTTCAGCGTT	7195	
Db	5520		GAGGAGATATCGTTCGCGCGGAGATCTTCGAAATCCAGGAAGTTCCTTCAGCGTT	5579	
Qy	7196		GCCATATGGGACGCGCGGACTACAAATCTCCACTCTGTAGAGTCTGGAAGACCGGA	7255	
Db	5580		GCCATATGGGACGCGCGGACTACAAATCTCCACTCTGTAGAGTCTGGAAGACCGGA	5639	
Qy	7256		CTAGTCCCTCCGGTGTACAGGATGCCATTTGCCACCTACCAAGGCTCTCCATACC	7315	
Db	5640		CTAGTCCCTCCGGTGTACAGGATGCCATTTGCCACCTACCAAGGCTCTCCATACC	5699	
Qy	7316		ACCTCCACGAGAAAGGACGGTTGCTCGACAGAAATCAATGTGCTTCTGCTTGGC	7375	
Db	5700		ACCTCCACGAGAAAGGACGGTTGCTCGACAGAAATCAATGTGCTTCTGCTTGGC	5759	
Qy	7376		GGAGCTCGCCACTAAGACCTTCGGTAGCTCCGGATCGTCGGCGTGTGATAGCGCACGGC	7435	
Db	5760		GGAGCTCGCCACTAAGACCTTCGGTAGCTCCGGATCGTCGGCGTGTGATAGCGCACGGC	5819	
Qy	7436		GACCGCCCTTCTGACCTGSCCTCCGACGAGTGACAAAGGATCGACGTTGAGTCGTA	7495	
Db	5820		GACCGCCCTTCTGACCTGSCCTCCGACGAGTGACAAAGGATCGACGTTGAGTCGTA	5879	
Qy	7496		CTCTCCATGCCCCCTTGAAGGGAGCCGGGGACCCCGATCTCAGGACGGGTCTTG	7555	
Db	5880		CTCTCCATGCCCCCTTGAAGGGAGCCGGGGACCCCGATCTCAGGACGGGTCTTG	5939	
Qy	7556		GTCTACCGTGAGGAGGCTAGTAGGATGTCGTCTGCTGCTCAATGTCTATACGTG	7615	
Db	5940		GTCTACCGTGAGGAGGCTAGTAGGATGTCGTCTGCTGCTCAATGTCTATACGTG	5999	
Qy	7616		GACAGGCGCCCTGATCAGCCATCGCTGCGGAGGAAAGTAAGTGCCTCAATCAACCGTT	7675	
Db	6000		GACAGGCGCCCTGATCAGCCATCGCTGCGGAGGAAAGTAAGTGCCTCAATCAACCGTT	6059	
Qy	7676		GAGCAACTTTTGTGCTGCTACCAACATGCTTACGCCAACATCCCGCAGCGCAAG	7735	
Db	6060		GAGCAACTTTTGTGCTGCTACCAACATGCTTACGCCAACATCCCGCAGCGCAAG	6119	
Qy	7736		CCTCGGACAGAGAGGTCACTTTGACAGATTTGCAAGTCTCGATGATCATTAACCGGA	7795	
Db	6120		CCTCGGACAGAGAGGTCACTTTGACAGATTTGCAAGTCTCGATGATCATTAACCGGA	6179	
Qy	7796		CGTACTCAAGAGATGAAGCGGAGCGTCCACAGTTAAGGCTAAGCTTCTATATAGA	7855	
Db	6180		CGTACTCAAGAGATGAAGCGGAGCGTCCACAGTTAAGGCTAAGCTTCTATATAGA	6239	
Qy	7856		GGAGGCTGCAAGCTGACGCCCCCACTTCGGCCAAATCCAAATTTGGCTATGGGCAAA	7915	

Db	6240		GGAGGCTGCAAGCTGACGCCCCACATTCGGCCAAATCCAAATTTGGCTATGGGCAAA	6299	
Qy	7916		GGAGCTCCGGAACCTATTCAGCAGGCGGCTTTAAACACATCCGCTCGTGTGGAGACTTT	7975	
Db	6300		GGAGCTCCGGAACCTATTCAGCAGGCGGCTTTAAACACATCCGCTCGTGTGGAGACTTT	6359	
Qy	7976		GCTGAGACACTGAAACACCAATTTGACACCACTCATGGCAAAAGTGAGGTTTCTG	8035	
Db	6360		GCTGAGACACTGAAACACCAATTTGACACCACTCATGGCAAAAGTGAGGTTTCTG	6419	
Qy	8036		CGTCCAAACAGAGAGGAGGCGCAAGCAGCTCGCTTATCGTATTTCCAGACCTGGG	8095	
Db	6420		CGTCCAAACAGAGAGGAGGCGCAAGCAGCTCGCTTATCGTATTTCCAGACCTGGG	6479	
Qy	8096		AGTTCGTGTATCGGAGAGATGGCCTTTTACGAGTGTCTCCACCTTCTCAGCGCT	8155	
Db	6480		AGTTCGTGTATCGGAGAGATGGCCTTTTACGAGTGTCTCCACCTTCTCAGCGCT	6539	
Qy	8156		GATGGGCTCTCATACGGATTTCAATACCTCCCCCAAGCAGCGGTTCGAGTTCCTGTGAA	8215	
Db	6540		GATGGGCTCTCATACGGATTTCAATACCTCCCCCAAGCAGCGGTTCGAGTTCCTGTGAA	6599	
Qy	8216		TACCTGGAATCAAGAAATGCCCTATGGCTTCTCATATGACACCCGCTGTTTGAATC	8275	
Db	6600		TACCTGGAATCAAGAAATGCCCTATGGCTTCTCATATGACACCCGCTGTTTGAATC	6659	
Qy	8276		AACGGTCACTGAGAGTGACATTCGTGTGAGGAGTCAATTTTACCAATGTTGACTTGGC	8335	
Db	6660		AACGGTCACTGAGAGTGACATTCGTGTGTGAGGAGTCAATTTTACCAATGTTGACTTGGC	6719	
Qy	8336		CCCCGAGGCGCAGACAGGCCATAAGTCTGCTCAGAGCGGCTTTTATCGGGGGTCCCT	8395	
Db	6720		CCCCGAGGCGCAGACAGGCCATAAGTCTGCTCAGAGCGGCTTTTATCGGGGGTCCCT	6779	
Qy	8396		GACTAACTCAAAAGGCGAGAACTGCGGTTATCCCGGTGCGCGCAAGTGGCTGTGAC	8455	
Db	6780		GACTAACTCAAAAGGCGAGAACTGCGGTTATCCCGGTGCGCGCAAGTGGCTGTGAC	6839	
Qy	8456		GACTAGCTGGGTAAATACCTCACATGTTACTTGAAGGCCACTGACGCTGTGAGCTGC	8515	
Db	6840		GACTAGCTGGGTAAATACCTCACATGTTACTTGAAGGCCACTGACGCTGTGAGCTGC	6899	
Qy	8516		AAAGTCCAGGACTGCAAGATGCTGTGAACGGAGACGACCTTGTCTGTTATCTGTGAAAG	8575	
Db	6900		AAAGTCCAGGACTGCAAGATGCTGTGAACGGAGACGACCTTGTCTGTTATCTGTGAAAG	6959	
Qy	8576		CGCGGAAACCCAGAGGATGCGGCGCCTTACGAGCTTTCACGGAGGCTATGACTAGTA	8635	
Db	6960		CGCGGAAACCCAGAGGATGCGGCGCCTTACGAGCTTTCACGGAGGCTATGACTAGTA	7019	
Qy	8636		TTCCGCCCCCCCCGGGATCCGCCCAACAGAAATACGACTTGGAGCTGATAACATCATG	8695	
Db	7020		TTCCGCCCCCCCCGGGATCCGCCCAACAGAAATACGACTTGGAGCTGATAACATCATG	7079	
Qy	8696		TTCTTCAATGTGTAGTCCGCGCAGATGCTTGGCAAAAGGCTATACCTTCAACCCG	8755	
Db	7080		TTCTTCAATGTGTAGTCCGCGCAGATGCTTGGCAAAAGGCTATACCTTCAACCCG	7139	
Qy	8756		TGACCCCAACACCCCTTGCACGGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAA	8815	
Db	7140		TGACCCCAACACCCCTTGCACGGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAA	7199	
Qy	8816		CTCTTGGCTAGGCAATATCATATGATGCGCCCACTTATGGCAAGGATGATTTCTGAT	8875	
Db	7200		CTCTTGGCTAGGCAATATCATATGATGCGCCCACTTATGGCAAGGATGATTTCTGAT	7259	
Qy	8876		GACTCACTTTTCTTCACTCTTCTAGCTCAAGACAACTTGAAGGCTTGGATTTGTCA	8935	
Db	7260		GACTCACTTTTCTTCACTCTTCTAGCTCAAGACAACTTGAAGGCTTGGATTTGTCA	7319	
Qy	8936		GATCTACGGGCTTGTCTCATTTGAGCCACTTGAACCTTACATCATTTGACGACT	8995	
Db	7320		GATCTACGGGCTTGTCTCATTTGAGCCACTTGAACCTTACATCATTTGACGACT	7379	

Db	2940	AAAGCTGACAGGCTCGGACGTGAACGCTGTAGCATATTACCGGGGCTTGATGTGTCGCT	2999
Qy	4616	CATACCGCTATCGGAGACGCTGCTGTGCGCAACAGACGCTCTAATGACGGGTTTCAC	4675
Db	3000	CATACCGCTATCGGAGACGCTGCTGTGCGCAACAGACGCTCTAATGACGGGTTTCAC	3059
Qy	4676	CGCGGATTTGACTAGTGATGACATGCGCAATACATGTGTACCCGACAGTTCAG	4735
Db	3060	CGCGGATTTGACTAGTGATGACATGCGCAATACATGTGTACCCGACAGTTCAG	3119
Qy	4736	CTTGGATCCCACTTACCATTTGAGACGACGCGTCCCAAGACGCGTGTGCGGCTC	4795
Db	3120	CTTGGATCCCACTTACCATTTGAGACGACGCGTCCCAAGACGCGTGTGCGGCTC	3179
Qy	4796	GCAACGGCGAGGTAGAACTGGCAGGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGG	4855
Db	3180	GCAACGGCGAGGTAGAACTGGCAGGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGG	3239
Qy	4856	AGAACGGCCCTCGGGCATGTTGCAATTTCTCGGCTCTGTGTGAGTGTATGACCGGGCTG	4915
Db	3240	AGAACGGCCCTCGGGCATGTTGCAATTTCTCGGCTCTGTGTGAGTGTATGACCGGGCTG	3299
Qy	4916	TGCTTGGTATGAGCTACGCGCCGCTGAGACTCGGTTAGGTTGCGGGCTTACCTAATAC	4975
Db	3300	TGCTTGGTATGAGCTACGCGCCGCTGAGACTCGGTTAGGTTGCGGGCTTACCTAATAC	3359
Qy	4976	ACCAGGGTTCCCGCTGCGCAGGACCATCTGGAGTCTTGGGAGAGCGCTTCTACAGGCT	5035
Db	3360	ACCAGGGTTCCCGCTGCGCAGGACCATCTGGAGTCTTGGGAGAGCGCTTCTACAGGCT	3419
Qy	5036	CACCCACATAGATGCCCATCTTCTGTCCAGATTAACAGGCGAGAGCAACTTTCCTTA	5095
Db	3420	CACCCACATAGATGCCCATCTTCTGTCCAGATTAACAGGCGAGAGCAACTTTCCTTA	3479
Qy	5096	CCTGGTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACTCCATCGTGGGA	5155
Db	3480	CCTGGTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACTCCATCGTGGGA	3539
Qy	5156	CCAAACTGTGAAGTGTCTCATACGGCTGAAACCTTACACTGCAGGGGCAACACCCCTGCT	5215
Db	3540	CCAAACTGTGAAGTGTCTCATACGGCTGAAACCTTACACTGCAGGGGCAACACCCCTGCT	3599
Qy	5216	GTATAGGCTAGGAGCCGTCGCAATAGAGTTCATCTCACACACCCCATAAATACAT	5275
Db	3600	GTATAGGCTAGGAGCCGTCGCAATAGAGTTCATCTCACACACCCCATAAATACAT	3659
Qy	5276	CATGGCATGATGTCGGCTGACCTGGAGGTGCTCACTAGCACCTGGGTGCTGTAGGCGG	5335
Db	3660	CATGGCATGATGTCGGCTGACCTGGAGGTGCTCACTAGCACCTGGGTGCTGTAGGCGG	3719
Qy	5336	AGTCTTGCAGCTTGGCGCATACTGCTGACACAGGAGTGTGCTCATTTGTGGGCGAG	5395
Db	3720	AGTCTTGCAGCTTGGCGCATACTGCTGACACAGGAGTGTGCTCATTTGTGGGCGAG	3779
Qy	5396	GATCATCTTGTCCGGGAGCGAGTGTGTTCCCGACAGGGAAGTCTCTACAGGAGTT	5455
Db	3780	GATCATCTTGTCCGGGAGCGAGTGTGTTCCCGACAGGGAAGTCTCTACAGGAGTT	3839
Qy	5456	CGATGAGATGAAGTGTGCTCACAATTTCTTATATGAGCAGGGAATGACGCTCGC	5515
Db	3840	CGATGAGATGAAGTGTGCTCACAATTTCTTATATGAGCAGGGAATGACGCTCGC	3899
Qy	5516	CGAGCAATTCAGCAAGGCGCTCGGTTGTTGCAACGCGCCACCAAGCAAGCGGAGGC	5575
Db	3900	CGAGCAATTCAGCAAGGCGCTCGGTTGTTGCAACGCGCCACCAAGCAAGCGGAGGC	3959
Qy	5576	TGCTGCTCCGCTGTGGAGTCAAGTGGCGAGCCCTTGAGACCTTCTGGGCGAAGCACAT	5635
Db	3960	TGCTGCTCCGCTGTGGAGTCAAGTGGCGAGCCCTTGAGACCTTCTGGGCGAAGCACAT	4019
Qy	5636	GTGGAATTTATACGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGCGGAAACCC	5695

Db	4020	GTGGATTTTCATCGCGGAATACAGTACCTAGCAGCTTATCCACTCTCCCTGGAAACCC	4079
Qy	5696	CGCGATAGCATCATTTGATGGCAATTTACAGCTTCTATCACTAGCCGCTCACACCCAAAA	5755
Db	4080	CGCGATAGCATCATTTGATGGCAATTTACAGCTTCTATCACTAGCCGCTCACACCCAAAA	4139
Qy	5756	CACCTCTCTGTTTAAACATCTTGGGGGATGGGGTGGCTGCCCACTCGCTCTCCAGCGC	5815
Db	4140	CACCTCTCTGTTTAAACATCTTGGGGGATGGGGTGGCTGCCCACTCGCTCTCCAGCGC	4199
Qy	5816	TGCGTCAGCTTCTGTTGGCGCGCATCGCCGAGCGGTGTGGCAGCATAGGCTTTGG	5875
Db	4200	TGCGTCAGCTTCTGTTGGCGCGCATCGCCGAGCGGTGTGGCAGCATAGGCTTTGG	4259
Qy	5876	GAAAGTGTCTGTGGAATCTTGGCGGGCTATGGGGCAGGGGTAGCGGCGCACTCGTGGC	5935
Db	4260	GAAAGTGTCTGTGGAATCTTGGCGGGCTATGGGGCAGGGGTAGCGGCGCACTCGTGGC	4319
Qy	5936	CTTTAAGGTCATGAGCGGAGGTGCTCCACCGAGGACCTGGTCAACTTACTCCCTGC	5995
Db	4320	CTTTAAGGTCATGAGCGGAGGTGCTCCACCGAGGACCTGGTCAACTTACTCCCTGC	4379
Qy	5996	CATCTCTCTCTGTTGCGCTCGGGGTCTGTGTGCGCAGCAATACTTCGCTCGGCA	6055
Db	4380	CATCTCTCTCTGTTGCGCTCGGGGTCTGTGTGCGCAGCAATACTTCGCTCGGCA	4439
Qy	6056	CGTGGCCCGGAGAGGGGGCTGTGAGTGAATGAACCGGCTGATAGCGTTCCGCTTCGCG	6115
Db	4440	CGTGGCCCGGAGAGGGGGCTGTGAGTGAATGAACCGGCTGATAGCGTTCCGCTTCGCG	4499
Qy	6116	GGGTACACAGTCTCCCTACGCACTATGTCCCTGAGAGCGCTGACAGCATGTGCAC	6175
Db	4500	GGGTACACAGTCTCCCTACGCACTATGTCCCTGAGAGCGCTGACAGCATGTGCAC	4559
Qy	6176	TCAGATCTCTCTGAGCTTACCACATCACTCAACTGCTGAAGCGCTCCACAGTGATTA	6235
Db	4560	TCAGATCTCTCTGAGCTTACCACATCACTCAACTGCTGAAGCGCTCCACAGTGATTA	4619
Qy	6236	TGAGGACTGTCTACGCCATGCTCCGGCTGCTGAGGATGAAGGATGTTTGGATGGATATG	6295
Db	4620	TGAGGACTGTCTACGCCATGCTCCGGCTGCTGAGGATGAAGGATGTTTGGATGGATATG	4679
Qy	6296	CACGTTTGTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCTGCGCGGTTTACCGGG	6355
Db	4680	CACGTTTGTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCTGCGCGGTTTACCGGG	4739
Qy	6356	AGTCCCTTTCTGTATGCAACCGCGGTACAAAGGAGTCTGCGGGGGGACGCGCATCAT	6415
Db	4740	AGTCCCTTTCTGTATGCAACCGCGGTACAAAGGAGTCTGCGGGGGGACGCGCATCAT	4799
Qy	6416	GCAAAACCATCTGCCATGCGGAGCAGATGCGCGGACATGTCAAAAACCGTTCCATGAG	6475
Db	4800	GCAAAACCATCTGCCATGCGGAGCAGATGCGCGGACATGTCAAAAACCGTTCCATGAG	4859
Qy	6476	GATCGTAGGCTTAGAACCTGACGCAACACGTCGCGGAGGAGTCTCCCATCAACGATA	6535
Db	4860	GATCGTAGGCTTAGAACCTGACGCAACACGTCGCGGAGGAGTCTCCCATCAACGATA	4919
Qy	6536	CACACGGGACCTTGACACCTCCCGCGGCCCACTATTTCAGGGCGCTATGCGCGGGT	6595
Db	4920	CACACGGGACCTTGACACCTCCCGCGGCCCACTATTTCAGGGCGCTATGCGCGGGT	4979
Qy	6596	GGCTGTGAGGAGTACGTTGAGGTTACGCGTGTGGGGATTTTCCACTACGTGACGGGAT	6655
Db	4980	GGCTGTGAGGAGTACGTTGAGGTTACGCGTGTGGGGATTTTCCACTACGTGACGGGAT	5039
Qy	6656	GACCATGACAAAGTAAAGTTCGCGGCTTCCGCGCCCGGAAATTTTTCACGGAGGT	6715
Db	5040	GACCATGACAAAGTAAAGTTCGCGGCTTCCGCGCCCGGAAATTTTTCACGGAGGT	5099
Qy	6716	GGATGAGTGGGTGACAGGTCAGCTCCGGCGTGCAAACTCTTCTACGGGAGGAGGT	6775
Db	5100	GGATGAGTGGGTGACAGGTCAGCTCCGGCGTGCAAACTCTTCTACGGGAGGAGGT	5159

QY 6776 CACGTTCCAGGTCGGGCTCAACCAATACTTGGTCGGTCGAGCTCCCATCGAGCCCGA 6835
DB 5160 CACGTTCCAGGTCGGGCTCAACCAATACTTGGTCGGTCGAGCTCCCATCGAGCCCGA 5219
QY 6836 ACCGACGTAACAGTGTCTTACTTCCATGCTCACCGATCCCTCCCAATTACAGCAGAGAC 6895
DB 5220 ACCGACGTAACAGTGTCTTACTTCCATGCTCACCGATCCCTCCCAATTACAGCAGAGAC 5279
QY 6896 GGCTAAGCGTAGGCTGGCTAGAGGCTCTCCCCCTCTTTTAGCCAGCTCATCAGCTAGCCA 6955
DB 5280 GGCTAAGCGTAGGCTGGCTAGAGGCTCTCCCCCTCTTTTAGCCAGCTCATCAGCTATCCA 5339
QY 6956 GTTGTCTCGCCTCTTTGAAGGCGACATGCACTACCCACCATGACTCCC CGGAGCGCTGA 7015
DB 5340 GTTGTCTCGCCTCTTTGAAGGCGACATGCACTACCCACCATGACTCCC CGGAGCGCTGA 5399
QY 7016 CTTCTATCGAGGCCAACCTCTTGTGCGCGAGAGATGGCGGAAACATCACTCGCGTGG 7075
DB 5400 CTTCTATCGAGGCCAACCTCTTGTGCGCGAGAGATGGCGGAAACATCACTCGCGTGG 5459
QY 7076 GTACAGATAAAGGTAGTAATTTCTGGAATCTTTTGAACCGGCTTACCGCGAGGGGATGA 7135
DB 5460 GTACAGATAAAGGTAGTAATTTCTGGAATCTTTTGAACCGGCTTACCGCGAGGGGATGA 5519
QY 7136 GAGGAGATATCCGTCGCGCGGAGATCTCGGAAATCCAGGAAAGTTCCCTCAGCGCTT 7195
DB 5520 GAGGAGATATCCGTCGCGCGGAGATCTCGGAAATCCAGGAAAGTTCCCTCAGCGCTT 5579
QY 7196 GCGCATATGGGCACGCCGAGCTACAATCTCTCACTGTCTAGAGTCTCTGGAAGAACCCGGA 7255
DB 5580 GCGCATATGGGCACGCCGAGCTACAATCTCTCACTGTCTAGAGTCTCTGGAAGAACCCGGA 5639
QY 7256 CTAAGTCCCTCCGCTGGTACA CGGATGCCCATTTGCCACTTACCAAGGCTCTCTCAATACC 7315
DB 5640 CTAAGTCCCTCCGCTGGTACA CGGATGCCCATTTGCCACTTACCAAGGCTCTCTCAATACC 5699
QY 7316 ACCTCCAGGAAAGAGACGGTGTCTGACAGATCCCAATGTCTTCTGCGCTTGGC 7375
DB 5700 ACCTCCAGGAAAGAGACGGTGTCTGACAGATCCCAATGTCTTCTGCGCTTGGC 5759
QY 7376 GGAGTTCGCCACTTAAGACCTTCGGTAGCTCCGGATCGTCGGCGTTGTATAGCGGCACGCG 7435
DB 5760 GGAGTTCGCCACTTAAGACCTTCGGTAGCTCCGGATCGTCGGCGTTGTATAGCGGCACGCG 5819
QY 7436 GACCGCCCTTCCTGACCTCGGCTCCGACGCGATGCAAAAGATCCGAGTTGATCGTA 7495
DB 5820 GACCGCCCTTCCTGACCTCGGCTCCGACGCGATGCAAAAGATCCGAGTTGATCGTA 5879
QY 7496 CTCCTCCATGCCCCCTTTGAAGGGAGCGCGGGACCCCGATCTCAGCGCGGCTTGG 7555
DB 5880 CTCCTCCATGCCCCCTTTGAAGGGAGCGCGGGACCCCGATCTCAGCGCGGCTTGG 5939
QY 7556 GTCTACCGGTGAGTAGGAGGCTAGTAGGATGTCTGCTGCTCAATGTCTTATACGTTG 7615
DB 5940 GTCTACCGGTGAGTAGGAGGCTAGTAGGATGTCTGCTGCTCAATGTCTTATACGTTG 5999
QY 7616 GACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAAGTAAAGTGCCTCAATCAACCCGTT 7675
DB 6000 GACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAAGTAAAGTGCCTCAATCAACCCGTT 6059
QY 7676 GAGCAACTCTTGTGCTGCTCACCACCAATGTCTTACGCCCAACATCCCGCAGCGCAAG 7735
DB 6060 GAGCAACTCTTGTGCTGCTCACCACCAATGTCTTACGCCCAACATCCCGCAGCGCAAG 6119
QY 7736 CTTCCGGCAGAGAAGGCTCACCTTTGACAGATTGCAAGTCTCTGGATGATCAATACCGGGA 7795
DB 6120 CTTCCGGCAGAGAAGGCTCACCTTTGACAGATTGCAAGTCTCTGGATGATCAATACCGGGA 6179
QY 7796 CGTACTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGGCTTCTATCTATAGA 7855
DB 6180 CGTACTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGGCTTCTATCTATAGA 6239

QY 7856 GGAGGCTCGAAGCTGACGCCCCCAACATTCGGCCAAATCCAAATTTGGCTATGGGGCAAA 7915
DB 6240 GGAGGCTCGAAGCTGACGCCCCCAACATTCGGCCAAATCCAAATTTGGCTATGGGGCAAA 6299
QY 7916 GGAGGCTCGGAAACCTATCCAGCAGGGCGGTTAACCAATCGCTCCGTTGGTGGAGGACTT 7975
DB 6300 GGAGGCTCGGAAACCTATCCAGCAGGGCGGTTAACCAATCGCTCCGTTGGTGGAGGACTT 6359
QY 7976 GCTTGAAGACACTGAACACCAATTGACACCAATCATGCAATCATGCAAAAGTGAAGTTTCTG 8035
DB 6360 GCTTGAAGACACTGAACACCAATTGACACCAATCATGCAAAAGTGAAGTTTCTG 6419
QY 8036 CGTCCAAACCCAGAGAGGAGGCGCCAGCAGCTCGCTTATCGTATTTCCAGACCTGGG 8095
DB 6420 CGTCCAAACCCAGAGAGGAGGCGCCAGCAGCTCGCTTATCGTATTTCCAGATTTGGG 6479
QY 8096 AGTTGCTGTATGCGAGAAAGATGGCCCTTTACGAGTGTCTCCACCTTCTCTCAGGCGCT 8155
DB 6480 GGTTCGTGTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCTCTCCAGGCGCT 6539
QY 8156 GATGGGCTCTCATACCGATTTCAATACTCCCAACGAGCGGCTCGAGTTCTCTGTGAA 8215
DB 6540 GATGGGCTCTTCATACCGATTTCAATACTCTCTGAGACAGGGGTGAGTTCTCTGGTGA 6599
QY 8216 TACCTGAAATCAAAAGAAATGCCCTATGCGGCTTCTCATATGACACCCGCTGTTTGAATC 8275
DB 6600 TGCCTGGAAGCGAAGAAATGCCCTATGCGGCTTTCGCATATGACACCCGCTGTTTGAATC 6659
QY 8276 AACCGTCACTGAGAGTGAATTCGTGTGAGGAGTCAATTTACCAATGTTGTGATCTGGC 8335
DB 6660 AACCGTCACTGAGAAATGACATCCGTTGAGGAGTCAATTCACCAATGTTGTGATCTGGC 6719
QY 8336 CCCCGAGGCGACAGCGCCATAGGTCGCTCACAGAGGGCTTTACATCGGGGCTCCCT 8395
DB 6720 CCCCGAGGCGACAGCGCCATAGGTCGCTCACAGAGGGCTTTACATCGGGGCGCCCT 6779
QY 8396 GACTAACTCAAAAGGCGAGAACTCGCGTTATCGCGGTGCGCGCAAGTGCCTGCTGAC 8455
DB 6780 GACTAACTCAAAAGGCGAGAACTCGCGTTATCGCGGTGCGCGCAAGTGCCTGCTGAC 6839
QY 8456 GACTAGCTCGGTTAATACCTTCACATGTTACTTGAAGGCCACTGACGCTCTGAGCTGTC 8515
DB 6840 GACTAGCTCGGTTAATACCTTCACATGTTACTTGAAGGCCCTGCGGCTCTGAGCTGTC 6899
QY 8516 AAAGCTCCAGGACTGCAAGTGTCTGACGAGAGCGACTTCTGTTATCTGTGAAG 8575
DB 6900 GAAGCTCCAGGACTGCAAGTGTCTGACGAGAGCGACTTCTGTTATCTGTGAAG 6959
QY 8576 CGCGGAAACCCAGGAGGATCGCGGCGCTTACGAGCGCTTACGAGGCTATGACTAGTA 8635
DB 6960 CGCGGAGAACCCAGAGGAGCGGCTTACGCGGCTTACGAGGCTATGACTAGTA 7019
QY 8636 TTCCGCCCCCCCCCGGGATCCGCCCAACCAAGATACGACTGAGCTGATAACATCATG 8695
DB 7020 CTCTGCCCCCTCGGGACCGCCCAAAACCAAGATACGACTTGGAGTTGATAACATCATG 7079
QY 8696 TTCTCTCAATGTGTGCTGCGGCGAGTGCATCTGGCAAAAGGTATATCTCTCACCG 8755
DB 7080 CTCTCTCAATGTGTGCTGCGGCGAGTGCATCTGGCAAAAGGTATATCTCTCACCG 7139
QY 8756 TGACCCCAACCCCGCTTGCACCGGCTGCGTGGGAGACAGTACACACTTCAATCAA 8815
DB 7140 TGACCCCAACCCCGCTTGGCGGGCTGCGTGGGAGACAGTACACACTTCAATCAA 7199
QY 8816 CTCTTGGCTAGGCAATATCATGTATGCGCCCAACCTATGGGCAAGGATGTTCTGAT 8875
DB 7200 TTCTCTGGCTAGGCAATATCATGTATGCGCCCAACCTTGTGGGCAAGGATGTTCTGAT 7259
QY 8876 GACTCACTTTTCTTCCATCTTCTAGCTCAAGAGCACTTCAAAAGCCCTGGATGTC 8935
DB 7260 GACTCACTTTTCTTCCATCTTCTAGCTCAGGAAACACTTGAAGGCGCTTCTAGTGTCA 7319
QY 8936 GATCTACGGGCTTGTCTACTCTTACCTTGTAGCCACTTGACCTACCTCAGATCATTTGAACGACT 8995

Db 7320 GATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTTCAGATCATTTCAAGCACT 7379
Qy 8996 CCATGCTCTTAGCGCATTTACACTCCACAGTTTACTCTCCAGGTGAGATCAATAGGGTGGC 9055
Db 7380 CCATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGC 7439
Qy 9056 TTCATGCCCTCAGGAACCTTGGGGTACCACTTGGGAACCTGGAGACATCGGGCCAGAAG 9115
Db 7440 TTCAATGCCCTCAGGAACCTTGGGGTACCGCTTGGAGTCTGGAGACATCGGGCCAGAA 7499
Qy 9116 TGTCCGGCTAGACTAGTCCAGAGGGGGAGGGCCGACATTGTGGCAGATACCTCTT 9175
Db 7500 TGTCCGGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGGAAGTACCTCTT 7559
Qy 9176 TAACTGGGAGTAAGGACCAAGCTTAAACTCACTCCAATCCCGCCGCGCTCCAGCTGGA 9235
Db 7560 CAACTGGGAGTAAGGACCAAGCTCAACTCACTCCCAATCCCGCTGCCAGTTGA 7619
Qy 9236 CTTGCTGCTGCTGCTCGTCTGGTTACAGCGGGGAGACATATATCAAGCCTGTCTCG 9295
Db 7620 TTTATCCAGCTGCTTCTGCTGTTACAGCGGGGAGACATATATCAAGCCTGTCTCG 7679
Qy 9296 TGCCCAACCCGCTGCTTCCGTTGCTGCTACTCTCTACTTCTGAGGGTAGGCATTTA 9355
Db 7680 TGCCCAACCCGCTGCTTCAATGCTGCTTCACTTCTTCTGAGGGTAGGCATCTA 7739
Qy 9356 CTGCTTCCCAACCGATGAACGGGAGCTTAAACACTCCAGGCTTAAGCC----- 9405
Db 7740 TCTACTCCCCACCGATGAACGGGAGCTTAAACACTCCAGGCTTAAGCC----- 9405
Qy 9406 ATTCTCTGTTTCT 9465
Db 7800 TTTTCCCTTCT 7859
Qy 9466 TCTTTTTTCTTTTCTTTTCCCTTCTTAAATGGTGGCTCCATCTTAGCCCTAGTCACGG 9525
Db 7860 TTTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7919
Qy 9526 CTAGCTGTGAAGTCCGTCAGCGCATGACTGCAGAGAGTGTGATACCTGGCCTCTCTG 9585
Db 7920 CTAGCTGTGAAGTCCGTCAGCGCTTGACTGCAGAGAGTGTGATACCTGGCCTCTCTG 7979
Qy 9586 CAGATCATGT 9595
Db 7980 CAGATCAAGT 7989

RESULT 5
US-10-509-921-9
; Sequence 9, Application US/10509921
; Publication No. US20050250093A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Hepatitis C Virus Sub-Genomic Replicons
; FILE REFERENCE: P51335
; CURRENT APPLICATION NUMBER: US/10/509,921
; CURRENT FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: 60/369,685
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 7979
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thepolynucleotide sequence encodes sequences from
; OTHER INFORMATION: HCV J4(B/RI) Replicons
US-10-509-921-9
Query Match 59.8%; Score 5733.6; DB 6; Length 7979;
Best Local Similarity 95.5%; Pred. No. 0;

Matches 5901; Conservative 0; Mismatches 279; Indels 0; Gaps 0;
Qy 3416 CTTTGGCCCATCATCGGCTACTCCCAACAAAGCGGGCGGTACTTGGTTGCATCATCAC 3475
Db 1800 CATGGCGCTATTACGGCTACTCCCAACAGACGCGAGCGCTACTTGGCTGCATCATCAC 1859
Qy 3476 TAGCCTCAGAGCCGGGACAAGAACAGGTCCGAAGGGAGGTTCAAAGTGGTTTCTACCGC 3535
Db 1860 TAGCCTCAGAGCCGGGACAAGAACAGGTCCGAAGGGAGGTTCAAAGTGGTTTCTACCGC 1919
Qy 3536 AACACAATCTTTCTGGCGACCTTGCAATCAACGGCGTGTGCTGACCTGTCTACCATGGCGC 3595
Db 1920 AACACAATCTTTCTGGCGACCTTGCAATCAACGGCGTGTGCTGACCTGTCTATCATGGTGC 1979
Qy 3596 TGGCTCGAAGACCTTAGCGGCTCCAAAAGGTCCAATCACCCAAATGTACACCAATGTAGA 3655
Db 1980 CGGCTCAAAAGACCTTTGCCGGGCCCAAAAGGGCCCAATCACCCAAATGTACACCAATGTAGA 2039
Qy 3656 CTGGACCTCGTGGCTGGCAGCGGCCCGCGGGGCGCTCCATGACACCATGTGCAAGCTG 3715
Db 2040 CMAAGACCTCGTGGCTGGCAAGCGGCCCGCGGGGCGGTTCTTTCACACCATGTGCAAGCTG 2099
Qy 3716 TGGCAGCTCGGACCTTTACTTTGCTCAGAGACATGTGTATGTCTATCCGGTGGCGCGCG 3775
Db 2100 CGGCAGCTCGGACCTTTACTTTGCTCAGAGGATGCCGATGTCTATCCGGTGGCGCGCG 2159
Qy 3776 AGCGCAGCAGAGGGGAAGTCTACTCTCCCCAGCGCGGCTCTCTACTCTGAAAGGCTCCTC 3835
Db 2160 GGGCGCAGCAGAGGGGAGGCTACTCTCCCCAGCGCGGCTCTCTACTCTGAAAGGCTCTTC 2219
Qy 3836 GGGTGTCTCAATTTGCTTTGCGCTTCGGGGCAGCTGCTGGGGCTCTTCGGGCTGTGTGTG 3895
Db 2220 GGGCGGTGCACCTGCTCTGCCCCCTCGGGGCGACGCTGTGGGCATCTTTTCGGGCTGCCGTGTG 2279
Qy 3896 CACCCGGGGGCTCGGAAAGCGGTGGACTTTCATACCCGTTGAGTCTATGGAACACTACCAT 3955
Db 2280 CACCCGAGGGGTTCCGAAGCGGTGGACTTTTGTACCGCTCGAGTCTATGGAACACTAT 2339
Qy 3956 GCGGTCTCGGCTTTCACAGACAACTCAACCCCCCGGCTGTACCGGACAGATTCAGAT 4015
Db 2340 GCGGTCCCCGGTCTTCACGGACAACTCGTCCCCCTCCGGCGGTATCCCGCAGACATTCACAGGT 2399
Qy 4016 GGCACATCTCAGCTCTACTCGCAGCGGACAGACACAAAGTGC CGGCTGGTATGC 4075
Db 2400 GGCCCATCTACCGCCCTACTGCTAGCGGCAAGACACTAAGGTGCCGCTGGCTGTATGC 2459
Qy 4076 AGCCCAAGGTTACAAGGTGCTCGTCTGAAACCGCTCGTTGCCGCCACCTTAGGGTTGG 4135
Db 2460 AGCCCAAGGTTAAGGTGCTTGTCTGAAACCGCTCGCTCGCGCCACCTTAGGTTTCGG 2519
Qy 4136 GGGCTATATGTCCAAGGCAACCGGTATCGACCTTAACATCAGAACTGGGGTAAAGACCAT 4195
Db 2520 GGGCTATATGTCTAAGGCACATGGTATCGACCTTAAACATCAGAAACCGGGGTAAAGACCAT 2579
Qy 4196 TACCAAGGGCGGCTCCATTAGCTACTCCACTATGGCAAGTTCTTCCCGACGGTGGCTG 4255
Db 2580 CACCACGGGTGCCCGCATACGTACTCACCTATGGAAGTGTCTTCCGCGAGGTTGGT 2639
Qy 4256 TTCTGGGGCGCTTATGACATCATATATGTGATGAGTGCACCTCAACTGACTCGACTAC 4315
Db 2640 CTCTGGGGCGGCTTATGACATCATATATGTGATGAGTGCACCTCAACTGACTCGACCAT 2699
Qy 4316 CATCTTGGGCATCGGCAAGTCTTGGACCAAGCGGAGACGGCTGGAGCGCGGCTGTCTGT 4375
Db 2700 TATCTTGGGCATCGGCAAGTCTTGGACCAAGCGGAGACGGCTGGAGCGCGGCTGTCTGT 2759
Qy 4376 GCTCGCACCGCTACACTCTCGGGATCGGTTACCGTGCACACCCCAATATCGAGGAAT 4435
Db 2760 GCTCGCACCGCTACCGCTCGGGATCGGTCACCGTGCACATCCAAACATCGAGGAGGT 2819
Qy 4436 AGGCTGTCCCAACATCGAGAGATCCCTCTTATGGCAAGGCCATCCCATTTAGGCCAT 4495
Db 2820 GGCTCTGTCCAGCACTGGAGAAATCCCTCTTTATGGCAAGGCCATCCCATTCGAGACCAT 2879

QY 4496 CAAGGGGGAGGCATCTCATTTTCTGCCATTTCCAAAGAAATGTTGACGAGCTCGCGC 4555
Db 2880 CAAGGGGGAGGCACCTCATTTTCTGCCATTTCCAAAGAAATGTTGATGAGCTCGCGC 2939
QY 4556 AAAGCTGACAGGCTTCGGACTGAACGCTGTAGCATATTTACCGGGCCTTGATGTCCGT 4615
Db 2940 GAAGCTGTCCGGCTTCGGACTCAATGCTGTAGCATATTTACCGGGCCTTGATGTATCCGT 2999
QY 4616 CATACCGCTATCGAGAGCTGTGTGTGTGGCAACAGACGCTCTAATGACGGGTTTCAC 4675
Db 3000 CATACCAACTAGCGGAGAGCTCATTTGTGTAGCAACGGACGCTCTAATGACGGGCTTTAC 3059
QY 4676 CGGCGATTTGACTCAGTGATGCGACTGCAATATACATGTGTACCCAGACAGTCGACTTCAG 4735
Db 3060 CGGCGATTTGCGACTCAGTGATGCGACTGCAATATACATGTGTACCCAGACAGTCGACTTCAG 3119
QY 4736 CTTGGATCCCACTTCCACATTTGAGACGACGACCGTGCCCAAGACGGGGTGTGCGGCTC 4795
Db 3120 CTTGGACCGGACCTTCCACATTTGAGACGACGACCGTGCCCAAGACGGGGTGTGCGGCTC 3179
QY 4796 GCAACGGCGAGGTAGAACTGGCAGGGGTAGGATGGCATCTTACAGGTTTGTGATCTCCAGG 4855
Db 3180 GCAGCGCGAGCGAGCACTGGTAGGGGAGGATGGCATTTACAGGTTTGTGACTCCAGG 3239
QY 4856 AGAAGCGCCCTCGGCGATGTTGCAATTTCTCGGTCTGTGTGAGTGTATGACGGGGCTG 4915
Db 3240 AGAAGCGCCCTCGGCGATGTTGCAATTTCTCGGTCTGTGTGAGTGTATGACGGGGCTG 3299
QY 4916 TGTCTGGTATGAGCTCAGCGCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACCTAAATAC 4975
Db 3300 TGTCTGGTACGAGCTCAGCGCCGCGGAGACCTCAGTTAGTTGCGGGCTTACCTAAACAC 3359
QY 4976 ACCAGGTTGCCGCTGTGCGAGGACCAATCTGAGTTCTGGGAGAGGCTTTCACAGGCT 5035
Db 3360 ACCAGGTTGCCGCTGTGCGAGGACCAATCTGAGTTCTGGGAGAGGCTTTCACAGGCT 3419
QY 5036 CACCCACATAGATGCCACTTCTCTCCAGACTAAACAGGACGAGACACTTTCCTTA 5095
Db 3420 CACCCACATAGAGCGCCCATTTCTGTCCAGACTAAGCAGGACGAGGACAACTTCCCTTA 3479
QY 5096 CTTGTGGCATATCAAGCTTACAGTGTGCGCAGGCTCAAGCTCCACCTCCATCTGTGGGA 5155
Db 3480 CTTGTAGCATACAGGCTACGGTGTGCGCAGGCTCAGGCTCCACCTCCATCTGGGA 3539
QY 5156 CCAATGTGGAAAGTGTCTCATACGGCTGAAACCTACATGCAAGGCGCAACACCCCTGCT 5215
Db 3540 CCAATGTGGAAAGTGTCTCATACGGCTGAAAGCTACGCTGCAAGGCGCAACACCCCTGCT 3599
QY 5216 GTATAGGCTAGGACCGTCCAAATGAGTGCATCTCTCACACCCCATTAATACAT 5275
Db 3600 GTATAGGCTGGAGCGGTTCAAAACAGAGTTTACTACACACCCCATTAACAAATACAT 3659
QY 5276 CATGCGATGCATGTCCGGCTGACCTGGAGTCTGCTACCTAGCACCTGGGTCTGGTAGGGG 5335
Db 3660 CATGCGATGCATGTCCGGCTGACCTGGAGTCTGCTACGAGCACTGGGTCTGGTAGGGG 3719
QY 5336 AGTCCTTGCGATTTGGCCGCAATATGCTCTGACGACAGCAGTGTGTCATTTGGGCGAG 5395
Db 3720 AGTCCTTAGCAGCTCTGGCCGCGTATTGCTGCAACAGGACGCTGTTCAITTTGGGCGAG 3779
QY 5396 GATCATCTTTGCGGGAAGCGAGTGTCTGTTCCCGACAGGGAAGTCTCTACAGGAGTT 5455
Db 3780 GATCATCTTTGCGGGAAGCGGCGCATCATTTCCCGACAGGGAAGTCTTTACCGGGAGTT 3839
QY 5456 CGATGAGATGGAAGAGTGTGCTCACAACCTTCTTACATCGAGCGGGAATGACGCTCGC 5515
Db 3840 CGATGAGATGGAAGAGTGTGCTCACAACCTTCTTACATCGAACAGGGAATGACGCTCGC 3899
QY 5516 CGAGCAATTTCAAGCAAAAGCGCTCGGGTTGTGCAAAAGCGGCAACCAAGCGGAGGC 5575
Db 3900 CGAAACAAATTCAAACAGAAAGCAATCGGGTTGTGTGCAAAAGCGGCAACCAAGCGGAGGC 3959

QY 5576 TGCTGCTCCGTGTGTGAGTCCAAAGTGGCGAGCCCTTTGAGACCTTTGCGGGCAAGACAT 5635
Db 3960 TGCTGCTCCGTGTGTGGAATCCAAAGTGGCGAGCCCTCGAAGCCCTTTGCGGGCAAGACAT 4019
QY 5636 GTGGAATTTTCAATAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTCTGGAACCC 5695
Db 4020 GTGGAATTTTCAATAGCGGGATACAAATATTTAGCAGGCTTGTCCACTCTGCTCTGCAACC 4079
QY 5696 CGCGATAGCATCATTTGATGCGATTTTACAGCTTCTATCACTAGCCGCTCACACCCAAA 5755
Db 4080 CGCGATAGCATCATTTGATGCGATTTTACAGCTTCTATCACAGCCGCTCACACCCAAA 4139
QY 5756 CACCTCTCTGTTTAACTCTTTGGGGGATGGGTGGCTGCCAACTCGCTCTCTCCAGCGC 5815
Db 4140 TACCTCTCTGTTTAACTCTTTGGGGGATGGGTGGCTGCCAACTCTCTCTCTCCAGCGC 4199
QY 5816 TGCCTCAGCTTTCTGTGGGCGCGCATCGCCGAGCGGCTTTTGGCAGCATAGGCTTGG 5875
Db 4200 TGCCTCAGCTTTCTGTAGCGCGCGCATCGCTGGAGCGGCTTTTGGCAGCATAGGCTTGG 4259
QY 5876 GAAAGTGTCTGTGACATCTTTGGCGGGCTATGGGCGAGGGGTAGCCGCGCATCTGTCGC 5935
Db 4260 GAAAGTGTCTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCGTGGC 4319
QY 5936 CTTTAAAGTCAATAGCGCGAGGTGCCCTCCACGAGGACCTGTGTCAACTTACTCCCTGC 5995
Db 4320 CTTTAAAGTCAATAGCGCGGAGATGCCCTCCACGAGGACCTGGTTAACTTACTCCCTGC 4379
QY 5996 CATCTCTCTCTCGTGGTCCCTGTGCTGGGGTCTGTGTGCGCAGCAATATCTCGGTGGCA 6055
Db 4380 TATCTCTCTCCCTCGCGCCCTAGTCTGTGGGGTCTGTGTGCGCAGCATACTCGGTGGCA 4439
QY 6056 CGTGGGCGCGGAGAGGGGCTGTGCAGTGTATGAAACCGGCTGTAGGCTTCGCTTCGCG 6115
Db 4440 CGTGGGCGCGGAGGAGGGGCTGTGCAGTGTATGAAACCGGCTGTAGGCTTCGCTTCGCG 4499
QY 6116 GGGTAAACACAGCTCTCCCTACGCACTATGTGCTGAGAGCGACGCTGCAGCAGCTGTGCAC 6175
Db 4500 GGGTAAACACAGCTCTCCCTACGCACTATGTGCTGAGAGCGACGCTGCAGCAGCTGTGCAC 4559
QY 6176 TCAGATCTCTCTAGCTTTACCATCATCTCAAATGCTGTGAAGCGGCTCCACGATGATTA 6235
Db 4560 TCAGATCTCTCTAGCTTTACCATCATCTCAAATGCTGTGAAGCGGCTTCACGATGATTA 4619
QY 6236 TGAGGACTGTCTAGGCGATCTCGGCTCGTGGCTAAGGATGTTTGGGATGATG 6295
Db 4620 CGAGGACTGTCTCCAGCGCATCTCGGCTCGTGGCTAAGAGATGTTTGGGATGATG 4679
QY 6296 CACGCTGTGACTGACTTTCAAGACCTGGCTCCAGTCCAAACTCTCTGCGCGGTTTACC 6355
Db 4680 CACGCTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTCTGCGCGGTTTACC 4739
QY 6356 AGTCCTTTCTGTGATGCCAAACGGGGTAAAGGAGTCTGGCGGGGGAGCGGATCAT 6415
Db 4740 AGTCCTTTCTTCTCATGTCAAACGGTGTAAAGGAGTCTGGCGGGGGAGCGGATCAT 4799
QY 6416 GCAACACCTTCCCATCGGAGCAGCATCGCCGACATGTCMAAAGGTTTCCATGAG 6475
Db 4800 GCAACACCTTCCCATGTGAGCAGCATCACCGGACATGTGAAGGATGTTTCCATGAG 4859
QY 6476 GATCGTAGGGCTAGAACCTCGAGCAACCGTGGCAGCGAACGTTCCCATCAACGCGATA 6535
Db 4860 GATCGTAGGGCTAGAACCTGTATGAACGTTGGCATTGGACATTTCCCATTAACGCGTA 4919
QY 6536 CACCAAGGAGCTTGTGCACACCTTCCCGCGCCCAACTATTTCAGGGGCGCTATGCGGCT 6595
Db 4920 CACCAAGGAGCTTGTGCACACCTTCCCGCGCCCAAAATTTATTCTAGGGGCGCTGTGCGGCT 4979
QY 6596 GGCTGCTGAGGAGTACGTGGAGTTTACCGGCTGGGGGATTTCCACTAGCTCAGCGGCAT 6655
Db 4980 GGCTGCTGAGGAGTACGTGGAGTTTACCGGCTGGGGGATTTTCCACTAGCTCAGCGGCAT 5039
QY 6656 GACCACTGACAAACGTTAAAGTGCCCATGCGAGGTTCCGGGCCCCCGAAATTTCTTACGAGGCT 6715

Db	5040	GACCACTGACAACTAAAGTGCCTGTCAGGTTCCGGCCCCCGAATCTTTCACGAGGT	5099
Qy	6716	GGATGAGTGCCTGTCACAGGTACGCTCCGGCGTCAAACTCTTCTTACGGGAGGACGT	6775
Db	5100	GGATGAGTGCCTGTCACAGGTACGCTCCGGCGTCAAACTCTTCTTACGGGAGGACGT	5159
Qy	6776	CACGTTCCAGGTTCGGGCTCAACCAATACTTTGGTTCGGGTCCGAGCTCCCATGCGAGCCGA	6835
Db	5160	CACGTTCCAGGTTCGGGCTCAACCAATACTTTGGTTCGGGTCCGAGCTCCCATGCGAGCCGA	5219
Qy	6836	ACCGGAGTAAAGTCTTACTTCCATGCTCACGATCCCTCCCAATACAGACAGAC	6895
Db	5220	ACCGGAGTAAAGTCTTACTTCCATGCTCACGATCCCTCCCAATACAGACAGAC	5279
Qy	6896	GGCTAAGCTGAGGTCTAGAGGTCTCCCTCCCTCTTTAGCCAGCTCATCAGCTAGCCA	6955
Db	5280	GGCTAAGCTGAGGTCTAGAGGTCTCCCTCCCTCTTTAGCCAGCTCATCAGCTATCCA	5339
Qy	6956	GTTGTCGCGCTTCTTTGAAGGCGACATGACATACCAATGACTCCCGGAGCGTGA	7015
Db	5340	GTTGTCGCGCTTCTTTGAAGGCGACATGACATACCAATGACTCCCGGAGCGTGA	5399
Qy	7016	CCTCATCGAGCCCACTCTTGTGGCGCAGGATGGGCGGAAACATCACTCGCGTGA	7075
Db	5400	CCTCATCGAGCCCACTCTTGTGGCGCAGGATGGGCGGAAACATCACTCGCGTGA	5459
Qy	7076	GTGAGAAATAAGTACTTAATCTGGACTCTTTTCCGAACCGCTTCAACGCGAGGGGATGA	7135
Db	5460	GTGAGAAATAAGTACTTAATCTGGACTCTTTTCCGAACCGCTTCAACGCGAGGGGATGA	5519
Qy	7136	GAGGAGATATCCGTCGCGCGGAGATCTTTCGGAATAACAGGAAGTTCCCTCAGCGTT	7195
Db	5520	GAGGAGATATCCGTCGCGCGGAGATCTTTCGGAATAACAGGAAGTTCCCTCAGCGTT	5579
Qy	7196	GCCCATATGGGCGAGCCCGGACTCAATCTTCCACTCTAGAGTCTTGGAGGACCCGGA	7255
Db	5580	GCCCATATGGGCGAGCCCGGACTCAATCTTCCACTCTAGAGTCTTGGAGGACCCGGA	5639
Qy	7256	CTAGCTCCCTCCGCTGTGTACAGGATGCCATTCGCACTTACCAAGGCTCTCCCAATACC	7315
Db	5640	CTAGCTCCCTCCGCTGTGTACAGGATGCCATTCGCACTTACCAAGGCTCTCCCAATACC	5699
Qy	7316	ACCTCCAGGAGAAAGGAGCGGTGTCTGACAGAAATCCAATGTCTTCTGCTTGGC	7375
Db	5700	ACCTCCAGGAGAAAGGAGCGGTGTCTGACAGAAATCCAATGTCTTCTGCTTGGC	5759
Qy	7376	GGAGCTCGCCACTAAGACCTTTCGGTGTGCTCCGGATCGTGGCCGTTGATAGCGCAGGC	7435
Db	5760	GGAGCTCGCCACTAAGACCTTTCGGTGTGCTCCGGATCGTGGCCGTTGATAGCGCAGGC	5819
Qy	7436	GACCGCTTCTGACCTGCTCCGAGCGGTGACAAAGGATCCGACGTTGAGTGTGTA	7495
Db	5820	GACCGCTTCTGACCTGCTCCGAGCGGTGACAAAGGATCCGACGTTGAGTGTGTA	5879
Qy	7496	CTCTCCATGCCCCCTTGAAGGGAGCCGGGGAGCCCGGATCTCAGCGACGGGCTTTG	7555
Db	5880	CTCTCCATGCCCCCTTGAAGGGAGCCGGGGAGCCCGGATCTCAGCGACGGGCTTTG	5939
Qy	7556	GTCTACCGTGTGAGGAGGCTGTGAGGATGTGCTGTGCTCAATGTCTTATACGTG	7615
Db	5940	GTCTACCGTGTGAGGAGGCTGTGAGGATGTGCTGTGCTCAATGTCTTATACGTG	5999
Qy	7616	GACAGGCCCTGATCAGCCATGCGGTGCGGAGGAAATGAGTGGCCCATCAACCGGTT	7675
Db	6000	GACAGGCCCTGATCAGCCATGCGGTGCGGAGGAAATGAGTGGCCCATCAACCGGTT	6059
Qy	7676	GAGCAACTCTTTGCTGCGTCAACCAATGCTTACGCCCAACATCCCGCAGCGCAAG	7735
Db	6060	GAGCAACTCTTTGCTGCGTCAACCAATGCTTACGCCCAACATCCCGCAGCGCAAG	6119
Qy	7736	CCTCGGCGAAGAGGTCACTTTTGACAGATTTGCAAGTCTCTGGATGATATACCGGGA	7795

Db	6120	CCTCCGGCAGAGAGGTCACTTTGACAGATTTGCAAGTCTCTGGATGATCATTTACCGGGA	6179
Qy	7796	CGTACTCAAGAGATGAAGGCGAGGCGTCCACAGTTAAGGCTAAGCTTCTATCTATAGA	7855
Db	6180	CGTACTCAAGAGATGAAGGCGAGGCGTCCACAGTTAAGGCTAAGCTTCTATCTATAGA	6239
Qy	7856	GGAGGCTCAAGCTGACGCGCCCAATTCGGGCAAAATCCAAATTTGGCTATGGGCAAA	7915
Db	6240	GGAGGCTCAAGCTGACGCGCCCAATTCGGGCAAAATCCAAATTTGGCTATGGGCAAA	6299
Qy	7916	GGAGCTCCGGAACCTATCCAGCAGGCGCTTAAACCAATCCCGCTCCGCTGTGGAGGACTT	7975
Db	6300	GGAGCTCCGGAACCTATCCAGCAGGCGCTTAAACCAATCCCGCTCCGCTGTGGAGGACTT	6359
Qy	7976	GCTGGAAGACACTGAAACCAATTCACACCAATTCATGGCAAAAGTAGGTTCTG	8035
Db	6360	GCTGGAAGACACTGAAACCAATTCACACCAATTCATGGCAAAAGTAGGTTCTG	6419
Qy	8036	CGTCCAAACAGAGAGGAGCGCGCAAGCCAGCTCCGCTTATCGTATTTCCAGACCTGGG	8095
Db	6420	CGTCCAAACAGAGAGGAGCGCGCAAGCCAGCTCCGCTTATCGTATTTCCAGACCTGGG	6479
Qy	8096	AGTTCTGTATGCGAGAAGATGCGCCCTTTACGACGCTGCTCCACCTTCTCAGGCCGT	8155
Db	6480	AGTTCTGTATGCGAGAAGATGCGCCCTTTACGACGCTGCTCCACCTTCTCAGGCCGT	6539
Qy	8156	GATGGGCTCTCATACGATTTCAATATCTCCCAAGCAGCGGTGAGTTCTGTGTAA	8215
Db	6540	GATGGGCTCTCATACGATTTCAATATCTCCCAAGCAGCGGTGAGTTCTGTGTAA	6599
Qy	8216	TACTCGAATCAAGAAATGCGCTTATGCGCTTCTCATATGACACCGCTGTTTGAATC	8275
Db	6600	TACTCGAATCAAGAAATGCGCTTATGCGCTTCTCATATGACACCGCTGTTTGAATC	6659
Qy	8276	AACGCTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACCAATGTTGTGCTTGGC	8335
Db	6660	AACGCTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACCAATGTTGTGCTTGGC	6719
Qy	8336	CCCCGAGGCGCAGACAGGCCATAAGTGTCTCAGAGCGGCTTTATATCGGGGCTCCCT	8395
Db	6720	CCCCGAGGCGCAGACAGGCCATAAGTGTCTCAGAGCGGCTTTATATCGGGGCTCCCT	6779
Qy	8396	GACTAATCAAAAGGCGAGAACTGCGGTATTCGCGGTGCGCGCAAGTGCCTGCTGAC	8455
Db	6780	GACTAATCAAAAGGCGAGAACTGCGGTATTCGCGGTGCGCGCAAGTGCCTGCTGAC	6839
Qy	8456	GACTAGCTGCGTAAATACCTCTCATGTTACTTGAAGGCCACTGACGCTGTCCAGCTGC	8515
Db	6840	GACTAGCTGCGTAAATACCTCTCATGTTACTTGAAGGCCACTGACGCTGTCCAGCTGC	6899
Qy	8516	AAAGCTCCAGGACTGCAAGTGTCTGTAAACGAGACGACCTTGTCTGTAAAG	8575
Db	6900	AAAGCTCCAGGACTGCAAGTGTCTGTAAACGAGACGACCTTGTCTGTAAAG	6959
Qy	8576	CGCGGAAACCGAGAGATGCGGGCGCTTACGAGCTTACGAGGCTATGACTAGTA	8635
Db	6960	CGCGGAAACCGAGAGATGCGGGCGCTTACGAGCTTACGAGGCTATGACTAGTA	7019
Qy	8636	TTCCGCGCGCGCGGATCCGCCCAACAGATAACGCTGAGCTGATAACATCATG	8695
Db	7020	TTCCGCGCGCGCGGATCCGCCCAACAGATAACGCTGAGCTGATAACATCATG	7079
Qy	8696	TTCTTCCAAATGTGTGCTGCGCAGATGCTATGCGCAAAAGGCTATACCTCACCCG	8755
Db	7080	TTCTTCCAAATGTGTGCTGCGCAGATGCTATGCGCAAAAGGCTATACCTCACCCG	7139
Qy	8756	TGACCCCAACCAACCCCTTGTGCAAGGCTGCGGTGGAGACAGCTAGACACCTCAATCAA	8815
Db	7140	TGACCCCAACCAACCCCTTGTGCAAGGCTGCGGTGGAGACAGCTAGACACCTCAATCAA	7199
Qy	8816	CTCTTGGCTAGGCAATCATCATGTATGCGCCCAACCTTATGGCAAGGATGTTCTCAT	8875
Db	7200	CTCTTGGCTAGGCAATCATCATGTATGCGCCCAACCTTATGGCAAGGATGTTCTCAT	7259

QY 8876 GACTCATTCTTCTCCATCCTTTCTAGCTCAAGAGCAACTTGAAAAAGCCCTGGATTGTCA 8935
DB 7260 GACTCATTCTTCTCCATCCTTTCTAGCTCAAGAGCAACTTGAAAAAGCCCTGGATTGTCA 7319
QY 8936 GATCTACGGGGTCTGCTACTCATTGAGCGCACTTGACCTCACTCAGATCATTTGAACGACT 8995
DB 7320 GATCTACGGGGTCTGCTACTCATTGAGCGCACTTGACCTCACTCAGATCATTTGAACGACT 7379
QY 8996 CCATGGTCTTAGCGCATTTACACTCCACAGTTTACTCTCCAGTGAGATCAATAGGGTGGC 9055
DB 7380 CCATGGTCTTAGCGCATTTACACTCCACAGTTTACTCTCCAGTGAGATCAATAGGGTGGC 7439
QY 9056 TTCTATGCCCTCAGGAAACTTGGGGTACCACCCCTTGCGAACTCGGAGACATCGGGCCAGAG 9115
DB 7440 TTCTATGCCCTCAGGAAACTTGGGGTACCACCCCTTGCGAACTCGGAGACATCGGGCCAGAG 7499
QY 9116 TGTCCGGCTTAAGCTACTGTCTCCAGGGGGAGGCGCGCACTTGTGGGAGATCACTCTT 9175
DB 7500 TGTCCGGCTTAAGCTACTGTCTCCAGGGGGAGGCGCGCACTTGTGGGAGATCACTCTT 7559
QY 9176 TAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAATCCGCGCGGCTCCAGCTGGA 9235
DB 7560 TAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCCAATCCGCGCGGCTCCAGCTGGA 7619
QY 9236 CTGTCTGGCTGGTTCTGGTCTGTTTACAGCGGGGGAGACATATACAGGCTGTCTCG 9295
DB 7620 CTGTCTGGCTGGTTCTGGTCTGTTTACAGCGGGGGAGACATATACAGGCTGTCTCG 7679
QY 9296 TGCCGAGCCCGCTCGGTTTCGTTGTGCTACTCTTCTAGGGGTAGGCAATTA 9355
DB 7680 TGCCGAGCCCGCTCGGTTTCGTTGTGCTACTCTTCTAGGGGTAGGCAATTA 7739
QY 9356 CTGTCTCCCAACCGATGACGGGAGCTTAACTCACTCCAGGCTTAAAGCCATTTCTCTTT 9415
DB 7740 CTGTCTCCCAACCGATGACGGGAGCTTAACTCACTCCAGGCTTAAAGCCATTTCTCTTT 7799
QY 9416 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 9475
DB 7800 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7859
QY 9476 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 9535
DB 7860 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7919
QY 9536 AAGGTCCTGAGCCGATGACTGCAGAGTCTGATCTGAGATCATGT 9595
DB 7920 AAGGTCCTGAGCCGATGACTGCAGAGTCTGATCTGAGATCAAGT 7979

RESULT 6
US-10-509-921-10
; Sequence 10, Application US/10509921
; Publication No. US20050250093A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Hepatitis C Virus Sub-Genomic Replicons
; FILE REFERENCE: P51335
; CURRENT APPLICATION NUMBER: US/10/509,921
; CURRENT FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: 60/369,685
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 7979
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Thepolynucleotide sequence encodes sequences from
; OTHER INFORMATION: HCV J4 (J4B/R1(C))Replicons
US-10-509-921-10

Query Match 59.7%; Score 5732; DB 6; Length 7979;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 5900; Conservative 0; Mismatches 280; Indels 0; Gaps 0;
QY 3416 CCTTGGCCCATCAGCGCTACTCCCAACAAACGCGGGCGTACTTGGTTGCATCATCAC 3475
DB 1800 CATGGCCCTATTATAGCCCTACTCCCAACAGACGAGGCGCTACTTGGCTGCATCATCAC 1859
QY 3476 TAGCCTCACAGGCGCGGACAAAGAACACAGGTTCGAAGGGAGGTTCAAGTGGTGTCTACCGC 3535
DB 1860 TAGCCTCACAGGCGCGGACAGGAACACAGGTTCGAAGGGAGGTCCAAGTGGTCTCACCGC 1919
QY 3536 AACAAATCTTTCTTGGGACCTGCATCAACGCGGTGTGTGGACTGTCTACCATGGCGC 3595
DB 1920 AACAAATCTTTCTTGGGACCTGCCTCAATGGGCTGTGTGGACTGTCTATCATGTGTG 1979
QY 3596 TGGCTCGAAGACCTTAGCGGTCCAAAAGGTCCAAATCACCACAAATGTACACCAATGTAGA 3655
DB 1980 CGGCTCAAGAGACCTTCCCGGCCCAAGGGGCCAATCACCACAAATGTACACCAATGTGA 2039
QY 3656 CTTGGACTCTCGTGGCTGGCAGGCGCCCCCGGGCGCGCTCCATGACACCATGCGAGCTG 3715
DB 2040 CCAGACCTCTGTGGCTGGCAGGCGCCCCCGGGCGCGGTCTCTTGACACCATGACACCTG 2099
QY 3716 TGGCAGCTCGGACCTTTTACTTGGTCAAGAGACATGTGATGTCAATCCGGTGGCGCGCG 3775
DB 2100 CGGAGCTCGGACCTTTTACTTGGTCAAGAGCATGCGGATGTCAATCCGGTGGCGCGCG 2159
QY 3776 AGGCGACAGAGGGAAGTCTACTCTCCCGAGCGCGCTCTCTACCTGAAAGGCTCTCTC 3835
DB 2160 GGGCGACAGCAGGGGAGCCTACTCTCCCGAGCGCGCTCTCTACTTTGAAAGGCTCTCTC 2219
QY 3836 GGGTGTCTCATTTGCTTGGCTTGGGCGAGCTCGTGGGCGCTCTTCCGGGCTGCTGTGTG 3895
DB 2220 GGGCGGTCTCACTGTCTTGGCTTGGGCGAGCTGTGGGCACTCTTCCGGGCTGCGGTGTG 2279
QY 3896 CACCGGGGGTCCGAGGCGGTGGACTTCATACCCGTTGAGTCTATGGAATACCAT 3955
DB 2280 CACCGAGGGGTTCGGAAGGCGGTGGACTTTGTACCCGTCGAGTCTATGGAACACCAT 2339
QY 3956 GCGGTCTCGGTCTTACAGACAACTCAACCCCGCGCGCTGTACGGGAGACATTTCCAAGT 4015
DB 2340 GCGGTCTCGGTCTTTCAGGACAACTCGTCCCTCCGCGCGTACGGGAGACATTTCCAAGT 2399
QY 4016 GGCACATCTGCAGCTCTCTACTGCGAGCGGAAGAGCACAAAGTCCGGCTGCTATGC 4075
DB 2400 GGCATCTACAGCGCTCTCTACTGAGCGGAGAGACAAAGTCCGGCTGCTATGC 2459
QY 4076 AGCCCAAGGGTACAGGTGTCTCTGAAACCCGTCGTTCCCGCCACTTATAGGTTTGG 4135
DB 2460 AGCCCAAGGGTATAAGGTGTCTCTGAAACCCGTCGTTCCCGCCACTTATAGGTTTGG 2519
QY 4136 GCGGTATATGTCGAAGGACAGGTATCGACCTTAACATCAGAACTGGGGTAAGGACCAT 4195
DB 2520 GCGGTATATGTCTAAGGCACATGGTATCGACCTTAACATCAGAACTGGGGTAAGGACCAT 2579
QY 4196 TACCACGGGCGCTCCATTAGCTACTCCACCTATGGAAGTTCCTTTCGCGCGGTGGGTG 4255
DB 2580 CACACAGGGTCCCCCATCAGTACTCCACCTATGGAAGTTCCTTTCGCGCGGTGGGTG 2639
QY 4256 TTCTGGGGCGCTATGACATCAATAATGTGAGTGGCCACTCAACTGACTCGACTAC 4315
DB 2640 CTCTGGGGCGCTATGACATCAATAATGTGAGTGGCCACTCAACTGACTCGACTAC 2699
QY 4316 CATCTTGGGCGATCGGACAGTCTTGGAACAGGAGACGGCTGGAGCGGCTCTGTCTGT 4375
DB 2700 TATCTTGGGCGATCGGACAGTCTTGGAACAGGAGACGGCTGGAGCGGCTCTGTCTGT 2759
QY 4376 GCTGCCACCGCTACACCTCTCGGGATCGGTTTACCGTGCACACCCCAATATCGAGGAAT 4435
DB 2760 GCTGCCACCGCTACCGCTCTCGGGATCGGTTTACCGTGCACATCCAAACATCGAGAGT 2819
QY 4436 AGGCTGTCCAACAAATGGAGAGATCCCTTCTATGGCAAGGCCATCCCAATTTGAGGCCAT 4495

Db 2820 GGCTCTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGCCAT 2879
Qy 4496 CAAGGGGGGAGGACATCTCATTTTCTGCCATTCGAAGAGAAATGTGACGAGCTCGCGG 4555
Db 2880 CAAGGGGGGAGGACATCTCATTTTCTGCCATTCGAAGAGAAATGTGAGAGCTCGCGG 2939
Qy 4556 AAAGCTGACAGGCTCGGACTGAACGCTGTAGATATTAACGGGGGCTTGTATGTCCGT 4615
Db 2940 GAAGCTGTCCGGCTCGGACTGAATGTGTAGCATATTAACGGGGGCTTGTATGTCCGT 2999
Qy 4616 CATACCGCCTATCGGAGAGCTGTGTCTGTGGCAACAGACGCTCTAATGACGGGTTTAC 4675
Db 3000 CATACCAACTAGCGGAGAGCTATTGTGTGAGCAACGAGCGCTCTAATGACGGGCTTAC 3059
Qy 4676 CGCGGATTTTGACTCAGTGTATCGACTCAATATACATGTGTCAACCCAGACAGTTCGACTTCAG 4735
Db 3060 CGGGGATTTGACTCAGTGTATCGACTCAATATACATGTGTCAACCCAGACAGTTCGACTTCAG 3119
Qy 4736 CTTGGATCCCACCTTCAACATTGAGACGACGACCGTGCCTCCCAAGACCGGCTGTGGGCTC 4795
Db 3120 CTTGGACCCGACCTTCAACATTGAGACGACGACCGTGCCTCCCAAGACCGGCTGTGGGCTC 3179
Qy 4796 GCAACGGCGAGGTAGAACTGCGACGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGG 4855
Db 3180 GCAGCGCGAGGCAAGGACTGTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGG 3239
Qy 4856 AGAACGGCCCTCGGGCATGTTGCAATTTCTGGTCTGTGTGAGTGTATGACGGGCTG 4915
Db 3240 AGAACGGCCCTCGGGCATGTTGCAATTTCTGGTCTGTGTGAGTGTATGACGGGCTG 3299
Qy 4916 TGCTTGGTATGAGCTACGCGCCGTGAGACCTCGGTTAGGTTGGGGCTTACCTAAATAC 4975
Db 3300 TGCTTGGTACGAGCTACGCGCCGCGAGACCTCAGTTAGGTTGGGGCTTACCTAAATAC 3359
Qy 4976 ACCAGGTTGGCCGTCTGCGAGGACCATCTGGAGTCTGGAGAGCGTCTTCAAGGCT 5035
Db 3360 ACCAGGTTGGCCGTCTGCGAGGACCATCTGGAGTCTGGAGAGCGTCTTCAAGGCT 3419
Qy 5036 CACCACATAGATGCCACTTCTGTCCGACACTAAACAGGACGAGACAACTTTCCTTA 5095
Db 3420 CACCACATAGATGCCACTTCTGTCCGACACTAAAGCAGGACGAGACAACTTTCCTTA 3479
Qy 5096 CCTGGTGATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCCATCGTGGGA 5155
Db 3480 CCTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCTCCATCGTGGGA 3539
Qy 5156 CCAATGTGGAAGTGTCTCATACGGCTGAACCTACACTGCAACGGGCGCAACCCCTGCT 5215
Db 3540 CCAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGCAACGGGCGCAACCCCTGCT 3599
Qy 5216 GTATAGGCTAGAGCCGTCCAAATGAGTGTATCTCATACACACCCCAATAAATACAT 5275
Db 3600 GTATAGGCTAGAGCCGTTCAAAACGAGGTTACTACCAACACCCCAATAAATACAT 3659
Qy 5276 CATGCACTGATGTCCGCTGACCTGGAGTGTCTCACTAGCACCTGGGTGCTGTAGGCGG 5335
Db 3660 CATGCACTGATGTCCGCTGACCTGGAGTGTCTCACTAGCACCTGGGTGCTGTAGGCGG 3719
Qy 5336 AGTCTTGCAGCTTTGGCGGCATACTGCTGTGACGACGAGGAGTGTGTCAATTGTGGGAG 5395
Db 3720 AGTCTTGCAGCTTTGGCGGCATTTGCTGTGACAAACGAGCGGTGTCAATTGTGGGAG 3779
Qy 5396 GATCATCTTGTCCGGGAAGCCAGCTGTGTGTCCGACAGGGAAGTCTCTACCGAGGTT 5455
Db 3780 GATCATCTTGTCCGGGAAGCCAGCCATCATTTCCGACAGGGAAGTCTCTTACCGGAGTT 3839
Qy 5456 CGATGAGTGAAGAGTGTGCTCACAACCTCTTACATCGACGAGGGAATGCACTCGC 5515
Db 3840 CGATGAGTGAAGAGTGTGCTCACAACCTCTTACATCGACGAGGGAATGCACTCGC 3899
Qy 5516 CGAGCAATTCAGGCAAAAGGCGCTCGGGTTGTGTGCAAAAGGCGCCACCAAGCAGCGGAGG 5575

Db 3900 CGAACAAATTCAAACAGAGGCAATCGGGTTGCTGCAAAACAGCCACCAAGCGAGGC 3959
Qy 5576 TGCTGCTCCGCTGGTGGAGTCCAAAGTGGCGAGCCCTTGAGACTTCTTGGCGGAGACAT 5635
Db 3960 TGCTGCTCCGCTGGTGGAAATCCAAAGTGGCGGACCTCTGAAAGCCTTCTGGCGGAAGATAT 4019
Qy 5636 GTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAACCC 5695
Db 4020 GTGGAATTTTCATCAGCGGATACATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCC 4079
Qy 5696 CGCGATAGCATCATTTGATGGCAATTTACAGCTTTATCATCTAGCCCCCTCACCAACCAAAA 5755
Db 4080 CGCGATAGCATCATCTGATGGCAATTTACAGCTTTATCATCCAGCCCTCATCAACAGCCCTCACCAACCAACA 4139
Qy 5756 CACCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCAACTCGCTCTCCAGCGC 5815
Db 4140 TACCTCTCTGTTTAAACATCTCGGGGGATGGGTGGCGGCCAACTTGTCTCTCCAGCGC 4199
Qy 5816 TGCGTCAGCTTTGCTGGCGCGCATCGCGAGCGGCTGTTGGCAGCATAGGCCCTTGG 5875
Db 4200 TGCTTCTGCTTCTGATAGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCCCTTGG 4259
Qy 5876 GAAGGTGCTCGTGGCAATCTTGGCGGGCTATGGGGCAGGGGTAGCCGGCGCATCTGCTGGC 5935
Db 4260 GAAGGTGCTGTTGATATTTTGGCAGTTATGAGCAGGGGTGGCAGGGCGCTCGTGGC 4319
Qy 5936 CTTTAAAGGTCAAGCGCGGAGGTGCCCTCAACGAGGACCTGGTCAACTTACTTCCCTGC 5995
Db 4320 CTTTAAAGGTCAAGCGCGGAGATGCCCTCACCGAGGACCTGGTTAACTTACTTCCCTGC 4379
Qy 5996 CATCTCTCTCTGCTGGCTGCTCGGGTCTGTCGGCAGCAATACTGCTCGGCA 6055
Db 4380 TATCTCTCTCCCTTGGCGCCCTAGTGTGCGGGTCTGTCGAGTGATGAACTGATGCTCGC 6115
Qy 6056 CTTGGGGCCCGGAGAGGGGCTGTGAGTGATGATGAACTGATGATGATGATGATGATGATG 6175
Db 4440 CTTGGGGCCCGGAGGGGCTGTGAGTGATGATGATGATGATGATGATGATGATGATG 4499
Qy 6116 GGGAACCAAGTCTTCCCTTACGCACTATGTCCTGAGAGGAGCGCTGCGAGCATGTCAC 6175
Db 4500 GGGTAAACAGCTCTCCCGCAGCACTATGTCCTGAGAGGAGCGCTGCGAGCATGTCAC 4559
Qy 6176 TCAGATCTCTTACGCTTACCATCTCACTCACTGCTGAAGCGCTCCACAGTGGATTA 6235
Db 4560 TCAGATCTCTTACGCTTACCATCTCACTCACTGCTGAAGCGCTTCCACAGTGGATTA 4619
Qy 6236 TGAGGACTGTCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACG 6295
Db 4620 CGAGGACTGTCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACG 4679
Qy 6296 CACGGTGTGACTGACTTCAAGACCTCGCTTCAAGTCCAACTCTGCGCGGGTTACCGGG 6355
Db 4680 CACGGTGTGACTGACTTCAAGACCTCGCTTCAAGTCCAACTCTGCGCGGGTTACCGGG 4739
Qy 6356 AGTCCCTTCTGTCATGCTCAACCGCGGTACAGGGAGTCTGCGGGGGAGCGGCATCAT 6415
Db 4740 AGTCCCTTCTGTCATGCTCAACCGCGGTACAGGGAGTCTGCGGGGGAGCGGCATCAT 4799
Qy 6416 GCRAACCACTGCGCCATGCGGAGCACAGATCGCGGACATGTCGCAAAACGGTTCATGAG 6475
Db 4800 GCAAAACCACTGCGCCATGCGGAGCACAGATCACCGGACATGTCGCAAAACGGTTCATGAG 4859
Qy 6476 GATCGTAGGGCTTAGAACCTTGCAGCAACACGTCGCGACGGAACGTTTCCCATTAACGCGTA 6535
Db 4860 GATCGTAGGGCTTAGAACCTTGCAGCAACACGTCGCGACGGAACGTTTCCCATTAACGCGTA 4919
Qy 6536 CACCAAGGACTTGTGCAACCTCCCGCGGCCCAACTATTTCCAGGGGCTTATGGCGGT 6595
Db 4920 CACCAAGGACTTGTGCAACCTCCCGCGGCCCAACTATTTCCAGGGGCTTATGGCGGT 4979
Qy 6596 GGCTGCTGAGGAGTACGTTGGAGGTTAGCGGTGTCGGGGATTTTCCACTACGTGACGGGAT 6655
Db 4980 GGCTGCTGAGGAGTACGTTGGAGGTTTACCGGGTGGGGATTTTCCACTACGTGACGGGAT 5039

QY 6656 GACCACTGACACGTAAGTGGCCATGCGAGTTCCGGCCCCCGAATCTTTCACGGAGGT 6715
DB 5040 GACCACTGACACGTAAGTGGCCGCTGTCAGGTTCCGGCCCCCGAATCTTTCACGGAGGT 5099
QY 6716 GGATGGAGTGGGTTGCAAGGTAAGCTCCGGCTGCAAACTCTTTCACGGAGGACGT 6775
DB 5100 GGATGGAGTGGGTTGCAAGGTAAGCTCCGGCTGCAAACTCTTTCACGGAGGACGT 5159
QY 6776 CACGTTCCAGTTCGGGCTCAACCAATACTTGGTCCGGTCCGAGCTCCCATGGAGCCCGA 6835
DB 5160 CACGTTCCAGTTCGGGCTCAACCAATACTTGGTCCGGTCCGAGCTCCCATGGAGCCCGA 5219
QY 6836 ACCGACGCTAACAGTGTCTTACTTCCATGCTACCCGATCCCTCCACATTTACAGCAGAGAC 6895
DB 5220 ACCGACGCTAACAGTGTCTTACTTCCATGCTACCCGATCCCTCCACATTTACAGCAGAGAC 5279
QY 6896 GGCTAAGCGTAGGCTGGCTAGAGGTCTCCCCCTCTTTAGCAGCTCATCAGCTAGCCA 6955
DB 5280 GGCTAAGCGTAGGCTGGCTAGAGGTCTCCCCCTCTTTAGCAGCTCATCAGCTATCCA 5339
QY 6956 GTTGTCTGGCCCTTCTTTGAAGGCGACATGCACTACCCACCATGACTCCCGGAGCGCTGA 7015
DB 5340 GTTGTCTGGCCCTTCTTTGAAGGCGACATGCACTACCCACCATGACTCCCGGAGCGCTGA 5399
QY 7016 CTTCTATCGAGGCCAACCTCTTGTGGCGGAGGAGATGGCGGAACATCACTTCGGTGGGA 7075
DB 5400 CTTCTATCGAGGCCAACCTCTTGTGGCGGAGGAGATGGCGGAACATCACTTCGGTGGGA 5459
QY 7076 GTCAGAGATAAGGTAGTAATTTCTGAGCTCTTTTCGAAACCGCTTTCACGCGGAGGGGATGA 7135
DB 5460 GTCAGAGATAAGGTAGTAATTTCTGAGCTCTTTTCGAAACCGCTTTCACGCGGAGGGGATGA 5519
QY 7136 GAGGAGATATCGGTGCGGGCGAGATCTCTGCGAATAATCCAGGAAGTTCCCTCAGCGTT 7195
DB 5520 GAGGAGATATCGGTGCGGGCGAGATCTCTGCGAATAATCCAGGAAGTTCCCTCAGCGTT 5579
QY 7196 GCCCATATGGGACGCCCGGACTACAACTCTCCACTGTCTGAGAGTCTCTGGAAGGACCCCGA 7255
DB 5580 GCCCATATGGGACGCCCGGACTACAACTCTCCACTGTCTGAGAGTCTCTGGAAGGACCCCGA 5639
QY 7256 CTACGTCCTCCGCTGGTACAGGATGCCCATTGGCCACCTTACCAAGGCTCTCTCCAATACC 7315
DB 5640 CTACGTCCTCCGCTGGTACAGGATGCCCATTGGCCACCTTACCAAGGCTCTCTCCAATACC 5699
QY 7316 ACCTCACGGAGAAAGAGACGGTTGTCTCTGACAGAAATCCAAATGTCTTCTGCTCTGGC 7375
DB 5700 ACCTCACGGAGAAAGAGACGGTTGTCTCTGACAGAAATCCAAATGTCTTCTGCTCTGGC 5759
QY 7376 GGAGCTCGGCCTTAAGACCTTCGGTAGCTCCGGATCGTCCGCGCTTGTATAGCGGCACCGC 7435
DB 5760 GGAGCTCGGCCTTAAGACCTTCGGTAGCTCCGGATCGTCCGCGCTTGTATAGCGGCACCGC 5819
QY 7436 GACCGCCCTTCTGACCTGGCTCCGACGACGGTGAACAAAGGATCCGAGTTGAGTTCGTA 7495
DB 5820 GACCGCCCTTCTGACCTGGCTCCGACGACGGTGAACAAAGGATCCGAGTTGAGTTCGTA 5879
QY 7496 CTCCTCATGCCCCCTTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTG 7555
DB 5880 CTCCTCATGCCCCCTTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTG 5939
QY 7556 GTCTACCGTAGTAGGAGGCTAGTCAGGATGTCTGCTCTGCTCAATGTCTTATACGTTG 7615
DB 5940 GTCTACCGTAGTAGGAGGCTAGTCAGGATGTCTGCTCTGCTCAATGTCTTATACGTTG 5999
QY 7616 GACAGCGCCCTGATCACGCCATGCTGCGGAGGAAAGTAAGCTGCCCATCAACCCGTT 7675
DB 6000 GACAGCGCCCTGATCACGCCATGCTGCGGAGGAAAGTAAGCTGCCCATCAACCCGTT 6059
QY 7676 GAGCAACTCTTGTCTGGTCAACCAAAATGTCTTAGCGGCAACAACTCCCGAGGCGCAAG 7735
DB 6060 GAGCAACTCTTGTCTGGTCAACCAAAATGTCTTAGCGGCAACAACTCCCGAGGCGCAAG 6119

QY 7736 CCTCCGCGACAGAAGGTCACTCTTTTGACAGATTGCAAGTCTCTGATGATCAATACCGGA 7795
DB 6120 CCTCCGCGACAGAAGGTCACTCTTTTGACAGATTGCAAGTCTCTGATGATCAATACCGGA 6179
QY 7796 CGTACTCAAGGAGATGAAGGGAAGGCGTCCACAGTTAAAGCTTAAGCTTCTATCTATAGA 7855
DB 6180 CGTACTCAAGGAGATGAAGGGAAGGCGTCCACAGTTAAAGCTTAAAGCTTCTATCTATAGA 6239
QY 7856 GGAGGCTGCAAGCTGACGCCCCACATTCGGCCAAATCCAAATTTGGCTATGGGCAAA 7915
DB 6240 GGAGGCTGCAAGCTGACGCCCCACATTCGGCCAAATCCAAATTTGGCTATGGGCAAA 6299
QY 7916 GGAGCTCCGGAACCTATCCAGCAGGCGGTTAAACCATCCGCTCCGTTGGGAGGACTT 7975
DB 6300 GGAGCTCCGGAACCTATCCAGCAGGCGGTTAAACCATCCGCTCCGTTGGGAGGACTT 6359
QY 7976 GCTGGAGACACTGTAACCAACCAATTTGACACCACTCATGSCAAAAGTGAAGTTTCTG 8035
DB 6360 GCTGGAGACACTGTAACCAACCAATTTGACACCACTCATGSCAAAAGTGAAGTTTCTG 6419
QY 8036 CGTCCAAACGAGAGAAGGAGGCGCAAGCCAGCTCGCTTATCTGTTATCCAGACCTCGG 8095
DB 6420 CGTCCAAACGAGAGAAGGAGGCGCAAGCCAGCTCGCTTATCTGTTATCCAGACCTCGG 6479
QY 8096 AGTTCTGTGTATGCGAGAAGATGGCCCTTTACGACGTGTCTTCCACCTTCTCTCAGGCCGT 8155
DB 6480 AGTTCTGTGTATGCGAGAAGATGGCCCTTTACGACGTGTCTTCCACCTTCTCTCAGGCCGT 6539
QY 8156 GATGGGCTCTCTATACGGATTTCAATACTTCCCCAAAGCAGCGGCTCGAGTTCTTGTA 8215
DB 6540 GATGGGCTCTCTATACGGATTTCAATACTTCCCCAAAGCAGCGGCTCGAGTTCTTGTA 6599
QY 8216 TACCTGGAATCAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTTGACTC 8275
DB 6600 TACCTGGAATCAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTTGACTC 6659
QY 8276 AACGCTCACTGAGAGTGACATTCGTGTTGAGAGTCAATTTACCAATGTTGTGACTTGGC 8335
DB 6660 AACGCTCACTGAGAGTGACATTCGTGTTGAGAGTCAATTTACCAATGTTGTGACTTGGC 6719
QY 8336 CCCGAGGCCAGACAGGCCATTAAGTCTGCTCACAGCGGCTTTACATCGGGGGTCCCT 8395
DB 6720 CCCGAGGCCAGACAGGCCATTAAGTCTGCTCACAGCGGCTTTACATCGGGGGTCCCT 6779
QY 8396 GACTTAACCTCAAAAGGCGAGACTCGCGTTATCGCGTCCGCGCAAGTGCCTGCAC 8455
DB 6780 GACTTAACCTCAAAAGGCGAGAACTCGCGTTATCGCGTCCGCGCAAGTGCCTGCAC 6839
QY 8456 GACTAGCTGCGGTAAATACCTTCACATGTTACTTGAAGGCCACTGCGAGCTGTGAGCTGC 8515
DB 6840 GACTAGCTGCGGTAAATACCTTCACATGTTACTTGAAGGCCACTGCGAGCTGTGAGCTGC 6899
QY 8516 AAAGCTCAGGACTGACGATGCTGTGAACGGAGACGACCTTGTGCTGTTATCTGTGAAG 8575
DB 6900 AAAGCTCAGGACTGACGATGCTGTGAACGGAGACGACCTTGTGCTGTTATCTGTGAAG 6959
QY 8576 CGCGGAAACCCAGGAGGATCGCGGCCCTTACGAGCTTTCAGGAGGCTATGACTAGGTA 8635
DB 6960 CGCGGAAACCCAGGAGGATCGCGGCCCTTACGAGCTTTCAGGAGGCTATGACTAGGTA 7019
QY 8636 TTTCGCCCCCCCCGGGATCCGCCCCCAACAGAAATACGACCTGGAGCTGATAACATCATG 8695
DB 7020 TTTCGCCCCCCCCGGGATCCGCCCCCAACAGAAATACGACCTGGAGCTGATAACATCATG 7079
QY 8696 TTCTCTCAATGTGTGTCAGTTCGCGCAGCATGATCTTGGCAAAAGGATATCTACTCCTCACCG 8755
DB 7080 TTCTCTCAATGTGTGTCAGTTCGCGCAGCATGATCTTGGCAAAAGGATATCTACTCCTCACCG 7139
QY 8756 TGACCCCAACCCCTTTCAGCGGCTCGTGGGAGACAGTACACACCTTCCCAATCAA 8815
DB 7140 TGACCCCAACCCCTTTCAGCGGCTCGTGGGAGACAGTACACACCTTCCCAATCAA 7199
QY 8816 CTCTTGGCTAGGCAATATCATGTATGCGCCCCACCTATGGGCAAGGATGATCTGTAT 8875

Qy	4436	AGGCTGTCCAACAATGGAGAGATCCCCTTCTATGGCAAGCCATCCCCATTGAGGCCAT	4495
Db	2820	GGCTCTGTCCAGCAC TGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCGAGACCAT	2879
Qy	4496	CAAGGGGGGAGGCAATCTCAATTTTCTGCCAATCCAAAGAAATGTGAAGAGCTCGCCGC	4555
Db	2880	CAAGGGGGGAGGCACCTCATTTTTCTGCCAATCCAAAGAAATGTGATGAGCTCGCCGC	2939
Qy	4556	AAAGCTCACAGGCTCGGATGCAACGCTGTAGCATATTACCGGGGCTTGATGTGCCGT	4615
Db	2940	GAAGCTGTCCGGCTCGGATCAATGTGTAGCATATTACCGGGGCTTGATGTATCCGT	2999
Qy	4616	CATACCCCTTATCGGAGACGTCGTTGCTGGCAACAGACGCTCTAATGACGGGTTTCAC	4675
Db	3000	CATACCAACTAGCGGAGACGTCATTGTCTGTAAGCAAGGACGCTCTAATGACGGGCTTAC	3059
Qy	4676	CGGCGATTTTGACTCAGTGATCGACTGCAATACATGTGTCAACAGACAGTGCACATTCA	4735
Db	3060	CGGCGATTTTGACTCAGTGATCGACTGCAATACATGTGTCAACAGACAGTGCACATTCA	3119
Qy	4736	CTTGGAATCCACCTTCAACATTGAGACGACGACGTCGCCCAAGACGGGTGTGCGCTC	4795
Db	3120	CCTGGACCCGACCTTCAACATTGAGACGACGACGTCGCCCAAGACGGGTGTGACGCTC	3179
Qy	4796	GCAACGCGAGGTAGAACTGCGAGGGTGTAGAGTGGCATCTACAGGTTTCTGACTCCAGG	4855
Db	3180	GCACGGGAGGACGACTGGTAGGGCAGGATGGGCATTTACAGTTTGTGACTCCAGG	3239
Qy	4856	AGAACGGCCCTCGGGCATTGTTTCGATTTCTCGTCTCTGTGTAGTGTATGACCGGGCTG	4915
Db	3240	AGAACGGCCCTCGGGCATTGTTTCGATTTCTCGTCTCTGTGTAGTGTATGACCGGGCTG	3299
Qy	4916	TGCTTGGTATGAGCTCAGCCCGCTGAGACCTCGGTTTAGTTGCGGGCTTACCTAAATAC	4975
Db	3300	TGCTTGGTATGAGCTCAGCCCGCTGAGACCTCAGTTAGTTGCGGGCTTACCTAAACAC	3359
Qy	4976	ACCAAGGTTGCCCTCTGCCAGGACCATCTGGAGTCTCTGGAGAGCGTCTTTCACAGSCCT	5035
Db	3360	ACCAAGGTTGCCCTCTGGCAGGACCATCTGGAGTCTCTGGAGAGCGTCTTTCACAGSCCT	3419
Qy	5036	CACCCACATAGATGCCACTTCTCTGCCAGACTTAAACAGGACGAGACAACTTTCCTTA	5095
Db	3420	CACCCACATAGACGCCACTTCTCTGCCAGACTTAAACAGGACGAGACAACTTCCCTTA	3479
Qy	5096	CCTGGTGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCAACCTCATGTGGGA	5155
Db	3480	CCTGGTAGCATACCAAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCATGTGGGA	3539
Qy	5156	CCAAATGTGGAAGTGTCTACACGGCTGAACCTACACTGCACGGGCCAACACCCCTGCT	5215
Db	3540	CCAAATGTGGAAGTGTCTACACGGCTGAACCTACAGCTGACGGGCCAACACCCCTGCT	3599
Qy	5216	GTATAGGCTAGGAGCCGTCCTCAAAATAGGTCATCCTCACACACCCCATAAATACAT	5275
Db	3600	GTATAGGCTAGGAGCCGTTCAAAACGAGTTTACTACACACACCCCATACCAAATACAT	3659
Qy	5276	CATGGCATGCATGTCCGCTGACCTGGAGGTGCTCACTAGCACCTTGGGTGTGTGAGCGG	5335
Db	3660	CATGGCATGCATGTCCGCTGACCTGGAGGTGCTCACTAGCACACCTTGGGTGTGTGAGCGG	3719
Qy	5336	AGTCTTTCAGCTTTGGCCGCATACCTGCTGACACAGGACGTCGTGCTATTGTTGGCAG	5395
Db	3720	AGTCTTTCAGCTCTGGCCCGGTATTGCTGACAAACAGGACGTCGTGCTATTGTTGGCAG	3779
Qy	5396	GATCATCTTTGTCGGGAAGCCAGCTGTCTGTTTCCCGACAGGGAAGTCTCTTACCAGGAGTT	5455
Db	3780	GATCATCTTTGTCGGGAAGCCAGCCATCAITTTCCCGACAGGGAAGTCTTACCAGGAGTT	3839
Qy	5456	CGATGAGATGAAGAGTGTGCTCAAACTTCTTTCATCATCGAGCAGGGAATGCAGCTCGC	5515
Db	3840	CGATGAGATGAAGAGTGTGCGCTCACACTCTCCCTTACATCATCGAAACAGGGAATGCAGCTCGC	3899

5516	CGAGCAATTCAAGCAAAAGCGCGCTCGGGTGTGTGCAAAAGCGGCCACCAAGCAAGCGGAGGC	5575
3900	CGAACAAATTCAAACAGAGGCAATACGGGTGTGTGCAAAAGCGGCCACCAAGCAAGCGGAGGC	3959
5576	TGCTGCTCCGTTGGTGGAGTCCAAAGTGGCGAGCCCTTTGAGACTTCTGGCGGAAGACACAT	5635
3960	TGCTGCTCCGTTGGTGGAAATCCAAGTGGCGGACCCCTCGAAGCCCTTCTGGCGGAAGCATAT	4019
5636	GTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACCTCTGCTTGGAACCC	5695
4020	GTGGAATTTTCATCAGCGGAATACAAATATTAGCAGGCTTGTCCACTCTGCTTGCAACCC	4079
5696	CGCGATAGCATCATGATGGCATTTACAGCTTCTATCATAGCCCGCTCACACCCAAAA	5755
4080	CGCGATAGCATCATGATGGCATTTACAGCCTCTATCACAGCCCGCTCACACCCAAACA	4139
5756	CACCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCCAACTCGCTCTCCACAGCC	5815
4140	TACCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCGGCCAACTTGTCTCTCCACAGCC	4199
5816	TGCGTCAGCTTTTCGTGGCGCGCGGCATCGCGGAGCGGCTGTGTGGCAGCATAGGCTCTGG	5875
4200	TGCTTCTGCTTTTCGTAGCGCGCGCATCGCTGAGCGGCTGTGTGGCAGCATAGGCTCTGG	4259
5876	GAAGTGCTCTGTTGGACATCTTGGCGGGCTATATGGGCGAGGGGTAGCGGGCGCATCTGTGCG	5935
4260	GAAGTGCTTTGCGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGGCGCTCGTGGCG	4319
5936	CTTTTAAGGTCATGAGCGGCGAGGTGCCCTCCACCGAGGACCTGGTCAACTTACTCCTCTGC	5995
4320	CTTTTAAGGTCATGAGCGGCGAGATGCCCCCTCCACCGAGGACCTGGTAACTACTCTCCTGC	4379
5996	CATCCTCTCTCTGTTGTCCTGTCTGTGCGGGTCTGTGTGGCAGCAATACTGCGTGGCA	6055
4380	TATCCTCTCTCCCTTGGCGGCCCTAGTGTGCGGGTCTGTGTGGCAGCAATACTGCGTGGCA	4439
6056	CGTGGGCGCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTTCGCG	6115
4440	CGTGGGCGCGAGGGGCGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTTCGCG	4499
6116	GGGTAAACACGCTCTCCCTTAACGACATATGTGCTTGAAGGAGCGCTGCAGCAGTGTCTAC	6175
4500	GGGTAAACACGCTCTCCCCCAACGACATATGTGCTTGAAGGAGCGCTGCAGCAGTGTCTAC	4559
6176	TCAGATCTCTAGCCTTACCATCACTCAACTGCTGAAGCGGCTCCACAGTGGATTAA	6235
4560	TCAGATCTCTAGTCTTACCATCACTCAAGTCTGCTGAAGGAGCTTCCACAGTGGATCAA	4619
6236	TGAGGACTGTCTACGCGCATGTCTCCGCTCGTGGCTAAGGATGTTTGGGATTTGGATATG	6295
4620	CGAGGACTGTCTCAAGCGCATGCTCCGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATG	4679
6296	CACGGTGTGACTGAATTTCAAGACCTTGGCTTCAGTCCAAACTCTCTGCGCGGTTTACCGGG	6355
4680	CACGGTGTGACTGAATTTCAAGACCTTGGCTTCAGTCCAAAGTCTCTGCGCGGATTTGCGCGG	4739
6356	AGTCCCTTTCTGTTCATGCCAAACGCGGGTACAAGGAGTCTGCGGGGGGACGCGCATCAT	6415
4740	AGTCCCTTTCTTCTCATGTCAAGTGGGTACAAGGGAGTCTGCGGGGGGACGCGCATCAT	4799
6416	GCAAAACCACTGCCCATCGGAGCACAGATCGCGGACATGTCAAAACCGGTTTCCATGAG	6475
4800	GCAAAACCACTGCCCATGTGGAGCACAGATCAACGGACATGTGAAACCGGTTTCCATGAG	4859
6476	GATCGTAGGGCTTAGAACCTTCGACGACAAACGTTGGCACCGGAAGTTCCTCCATTCACGCAATA	6535
4860	GATCGTAGGGCTTAGAACCTGTAGTAACACGTTGGCATGGAAACATTCCTCCATTTAACCGGTA	4919
6536	CACACGGGACCTTGACACACCTCCCGGCGGCCAACTATTTCAGGGGCGCTATGGCGGGT	6595
4920	CACACGGGCGCCCTTGACACGCGCCCTCCCGGCGGCCAAATTTATTTAGGGGCGCTGTGGCGGGT	4979
6596	GGCTGCTGAGGAGTACGTGGAGGTTACGCGTGTGGGGGATTTTCCACTACGTGACGGGCAT	6655

Db 4980 GGTCTGCTGAGAGTACGTGAGGTTTACGGGGTGGGGATTTCCACTACGTGACGGGCAT 5039
Qy 6656 GACCACGTACAAAGTAAAGTGCATGCCAGGTTTCGGGCCCCCGAAATTTCTTACGGAGGT 6715
Db 5040 GACACGTGACAAAGTAAAGTGCATGCCAGGTTTCGGGCCCCCGAAATTTCTTACAGAGT 5099
Qy 6716 GGATGGAGTCGGTTGACAGGTACGTCCGGCGTGCAGAACCTTCTTACGGAGAGCT 6775
Db 5100 GGATGGGTCGGTTGACAGGTACGTCCAGCGTGCAGAACCTTCTTACGGAGAGGT 5159
Qy 6776 CAGTTCCAGGTGGGCTCAACCAATACCTTGGTCCGGTCCAGCTCCCATGCCAGCCGA 6835
Db 5160 CACATTCCTGTCGGGCTCAATCAATACCTTGGTCCAGCTCCCATGCCAGCCGA 5219
Qy 6836 ACCGAGCTAACAGTGTACTTCCATGCTCACCGATCCCTCCACATACAGCAGAGAC 6895
Db 5220 ACCGAGCTAGAGTGTCTACTTCCATGCTCACCGATCCCTCCACATACAGCAGAGAC 5279
Qy 6896 GGCTAACGCTAGGCTGAGAGGTCTCCCTCTTTAGCCAGCTCATCAGCTAGCCA 6955
Db 5280 GGCTAACGCTAGGCTGAGAGGTCTCCCTCTTTAGCCAGCTCATCAGCTAGCCA 5339
Qy 6956 GTTGTCTGCGCTTTTGAAGGCGAATGCACTACCCACATGACTCCCGGAGCGTGA 7015
Db 5340 GCTGTCTGCGCTTTTGAAGGCGAATGCACTACCCACATGACTCCCGGAGCGTGA 5399
Qy 7016 CTTTCATGAGCCCACTTGTGGCGGAGAGTGGGCGGGAACATCACTCCGCTGGA 7075
Db 5400 CTTTCATGAGCCCACTTGTGGCGGAGAGTGGGCGGGAACATCACTCCGCTGGA 5459
Qy 7076 GTACAGAAATAGGTAGTAAATTCGACTCTTTTCGAAACCGCTTCAACGAGGAGGATGA 7135
Db 5460 GTACAGAAATAGGTAGTAAATTCGACTCTTTTCGAGCGCTTCAACGAGGAGGATGA 5519
Qy 7136 GAGGAGATATCCCTGCGCGGAGATCTTTCGAGAAATTCAGGAGTTTCCCTCAGCGTT 7195
Db 5520 GAGGAGATATCCCTGCGCGGAGATCTTTCGAGAGGTTCAGGAGAAATTCCTCAGCGAT 5579
Qy 7196 GCCATATGCGCAGCGCGGACTCAATCTCCACTCTGACTAGAGTCTCGGAGGACCCGGA 7255
Db 5580 GCCATATGCGCAGCGCGGACTCAATCTCCACTCTGACTAGAGTCTCGGAGGACCCGGA 5639
Qy 7256 CTACGTCCTCGGTGTTACAGGATGCCATTCACCTACCAAGGCTCTCCCAATACC 7315
Db 5640 CTACGTCCTCGGTGTTACAGGATGCCATTCACCTACCAAGGCTCTCCCAATACC 5699
Qy 7316 ACCTCCAGGAGAAAGGACGGTTGTTCTGACAGAAATTCGAATGTTCTGCTTGGC 7375
Db 5700 ACCTCCAGGAGAAAGGACGGTTGTTCTGACAGAAATTCGAATGTTCTGCTTGGC 5759
Qy 7376 GGAGCTGCCACTAAGACTTCGGTAGCTCCGGATCGTGGCGGTTGATAGCGGACCGC 7435
Db 5760 GGAGCTGCCACTAAGACTTCGGTAGCTCCGGATCGTGGCGGTTGATAGCGGACCGC 5819
Qy 7436 GACGCGCTTCTGACCTGGCTCCGACGAGGTGACAAAGATCCGAGCTTGAGTCGTA 7495
Db 5820 AACGCGCTTCTGACACGAGCTTCCGACGAGGACCGGGATCCGACGTTGAGTCGTA 5879
Qy 7496 TCTCTCATGCCCCCTTTGAAGGGGAGCGGGGACCGCGATCTCAGCAGCGGCTTTG 7555
Db 5880 TCTCTCATGCCCCCTTTGAAGGGGAGCGGGGATCCCGATCTCAGCAGCGGCTTTG 5939
Qy 7556 GTCTACCGTAGTAGGAGGTAGTAGAGTGTCTGCTGCTCAATGCTCTATACGTG 7615
Db 5940 GTCTACCGTAGTAGGAGGTAGTAGAGTGTCTGCTGCTCAATGCTCTATACGTG 5999
Qy 7616 GACGCGCTTCTGACGCTGCTGCGGAGGAAAGTAAAGTGGCCATCAACCCGTT 7675
Db 6000 GACGCGCTTCTGACGCTGCTGCGGAGGAAAGTAAAGTGGCCATCAACCCGTT 6059
Qy 7676 GAGCAACTCTTGTGCTGCTGACCAACATGCTGTACGCCCAACATCCCGAGGCAAG 7735

6060 GAGCAACTCTTGTCTGCTGACCAACATGTTTACGCCCAACATCCGCGAGCGCAAG 6119
7736 CTTCCGGCAGAAAGAGTCACTTTTGAAGATTGCAAGTCTCTGGATGATCATTTACCGGA 7795
6120 CTTCCGGCAGAAAGAGTCACTTTTGAAGATTGCAAGTCTCTGGATGATCATTTACCGGA 6179
7796 CGTACTCAAGGAGATGAAGCGAAGCGTCCACAGTTAAAGCTTAAGCTTCTATCTATAGA 7855
6180 CGTACTCAAGGAGATGAAGCGAAGCGTCCACAGTTAAAGCTTAAGCTTCTATCTATAGA 6239
7856 GAGGCGCTGCAAGCTGACGCCCCCACAATTCGGGCAAAATCCAAATTTGGCTATGGGCAAA 7915
6240 GAGGCGCTGCAAGCTGACGCCCCCACAATTCGGGCAAAATCCAAATTTGGCTATGGGCAAA 6299
7916 GAGCGTCCGGAACCTATCCAGAGGCGGTTTAAACATCCGCTTCGTTGGGAGGACTT 7975
6300 GAGCGTCCGGAACCTATCCAGAGGCGGTTTAAACATCCGCTTCGTTGGGAGGACTT 6359
7976 GCTGGAGACACTGAACCAACCAATTCACACCAATTCATGGCAAAAGTGAAGTTTCTG 8035
6360 GCTGGAGACACTGAACCAACCAATTCACACCAATTCATGGCAAAAGTGAAGTTTCTG 6419
8036 CGTCCAAACAGAGAGGAGCGGCGCAAGCAGCTCGCTTATCGTATTCAGACCTGGG 8095
6420 CGTCCAAACAGAGAGGAGCGGCGCAAGCAGCTCGCTTATCGTATTCAGACCTGGG 6479
8096 AGTTCGTGTATGCGAGAGATGCCCTTTACAGAGTGGTCTCAACCTTCTCAGGCGCT 8155
6480 AGTTCGTGTATGCGAGAGATGCCCTTTACAGAGTGGTCTCAACCTTCTCAGGCGCT 6539
8156 GATGGGCTCTCATAGGATTTCAATCTCCCCCAAGCAGCGGTCGAGTTCCTGGTGAA 8215
6540 GATGGGCTCTCATAGGATTTCAATCTCCCCCAAGCAGCGGTCGAGTTCCTGGTGAA 6599
8216 TACTCGAATCAAGAAATGCGCTTATGGCTTCTCATATGACACCGCTGTTTGAATC 8275
6600 TACTCGAATCAAGAAATGCGCTTATGGCTTCTCATATGACACCGCTGTTTGAATC 6659
8276 AACGGTCACTGAGAGTGACATTCGTGTGAGGAGTCAATTTACCAATGTTGTGACTGGC 8335
6660 AACGGTCACTGAGAGTGACATTCGTGTGAGGAGTCAATTTACCAATGTTGTGACTGGC 6719
8336 CCGGAGGCGCAGACAGGCCATAGGTGCTCAGAGCGGCTTTACATCGGGGTCCTT 8395
6720 CCGGAGGCGCAGACAGGCCATAGGTGCTCAGAGCGGCTTTACATCGGGGTCCTT 6779
8396 GACTTAACCTCAAAAGGCGAGAGTGGGCTTATCGCGGTGCGGCAAGTGGCGTGTGAC 8455
6780 GACTTAACCTCAAAAGGCGAGAGTGGGCTTATCGCGGTGCGGCAAGTGGCGTGTGAC 6839
8456 GACTTAGCTGCGGTAATACCTCTACATGTTACTTGAAGGCCACTGACGCTGTGAGCTGC 8515
6840 GACTTAGCTGCGGTAATACCTCTACATGTTACTTGAAGGCCACTGACGCTGTGAGCTGC 6899
8516 AAAGCTCCAGAGTGCAGATGCTGTGAACCGGAGACGACCTTGTGCTTATCTGTGAAG 8575
6900 AAAGCTCCAGAGTGCAGATGCTGTGAACCGGAGACGACCTTGTGCTTATCTGTGAAG 6959
8576 CGCGGGAACCGAGAGATGGGCGGCTTACGAGCTTTCAGCGAGGCTATGACTAGGTA 8635
6960 CGCGGGAACCGAGAGATGGGCGGCTTACGAGCTTTCAGCGAGGCTATGACTAGGTA 7019
8636 TTCCGCGCCCCCGGGGATCCGCCCAACAGAAATACGACTGGAGCTGATAACATCATG 8695
7020 TTCCGCGCCCCCGGGGATCCGCCCAACAGAAATACGACTGGAGCTGATAACATCATG 7079
8696 TTCTCTCAATGTGTGAGTGCAGTGCATTTGGCAAAAGGTAATACCTACCTCACCCG 8755
7080 TTCTCTCAATGTGTGAGTGCAGTGCAGTGCATTTGGCAAAAGGTAATACCTACCTCACCCG 7139
8756 TGACCCCAACACCCCTTGTGACCGGGTGTGCGGAGACAGCTAGACACACTCCCAATCAA 8815
7140 TGACCCCAACACCCCTTGTGACCGGGTGTGCGGAGACAGCTAGACACACTCCCAATCAA 7199

QY 8816 CTCCTGGCTAGGCAATATCATCATGATGCGCCACCCTATGGGCAAGGATGATCTTGAT 8875
DB 7200 CTCCTGGCTAGGCAATATCATCATGATGCGCCACCCTATGGGCAAGGATGATCTTGAT 7259
QY 8876 GACTCAGCTTTTCTCCATCTCTAGCTCAAGAGCACTTGAAAAGCCCTGGATGTCFA 8935
DB 7260 GACTCAGCTTTTCTCCATCTCTAGCTCAAGAGCACTTGAAAAGCCCTGGATGTCFA 7319
QY 8936 GATCTACGGGGCTTGCTACTCCATTGAGCCACTTGACCTACTCAGATCATTTGACGACT 8995
DB 7320 GATCTACGGGGCTTGCTACTCCATTGAGCCACTTGACCTACTCAGATCATTTGACGACT 7379
QY 8996 CCATGGCTCTTAGCGCATTTACACTCCACAGTTTACTCTCCAGTGTAGATCAATAGGGTGGC 9055
DB 7380 CCATGGCTCTTAGCGCATTTACACTCCACAGTTTACTCTCCAGTGTAGATCAATAGGGTGGC 7439
QY 9056 TTCATGGCTCAGGAACTTGGGGTACCACTCTTGCGAACTTGGAGACATCGGGCCAGAG 9115
DB 7440 TTCATGGCTCAGGAACTTGGGGTACCACTCTTGCGAACTTGGAGACATCGGGCCAGAG 7499
QY 9116 TGTCGCGCTAGCTACTGTCACGAGGGGAGGCGCCACTTGTGGCAGATACCTCTT 9175
DB 7500 TGTCGCGCTAGCTACTGTCACGAGGGGAGGCGCCACTTGTGGCAGATACCTCTT 7559
QY 9176 TAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCAATCCGCGCGGTGCCAGCTGGA 9235
DB 7560 TAACTGGGCGAGTAAGGACCAAGCTTAAACTCACTCAATCCGCGCGGTGCCAGCTGGA 7619
QY 9236 CTTGTCTGGCTGGTTGCTGCTGGTTTACAGCGGGGAGACATATATACAGCCTGTCTCG 9295
DB 7620 CTTGTCTGGCTGGTTGCTGCTGGTTTACAGCGGGGAGACATATATACAGCCTGTCTCG 7679
QY 9296 TGCCGACCCCGCTGGTTTCGTTGCTGCTACTCTTCTTCTGTAGGGTAGGCAATTA 9355
DB 7680 TGCCGACCCCGCTGGTTTCGTTGCTGCTACTCTTCTTCTGTAGGGTAGGCAATTA 7739
QY 9356 CTGCTGCCCAACCATGAACGGGAGCTTAACCACTCCAGGCTTTAAGCC----- 9405
DB 7740 CTGCTGCCCAACCATGAACGGGAGCTTAACCACTCCAGGCTTTAAGCC----- 7799
QY 9406 ATTTCCTGTTT 9465
DB 7800 TTTTCCCTT 7859
QY 9466 TCTTTTTTCTTTTCTTTTCCCTCTTTAATGGTGGCTCCATCTTAGCCCTAGTCAAGG 9525
DB 7860 TTTTTCCTTTTTTTTTTCTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7919
QY 9526 CTAGCTGTGAAGGTCGCGAGCGCATGACTGCGAGAGTCTGATCTAGTGGCTCTCTG 9585
DB 7920 CTAGCTGTGAAGGTCGCGAGCGCTTGACTGCGAGAGTCTGATCTAGTGGCTCTCTG 7979
QY 9586 CAGATCATCT 9595
DB 7980 CAGATCAACT 7989

RESULT 8
US-11-119-330-1
; Sequence 1, Application US/11119330
; Publication No. US20050260568A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Min
; APPLICANT: Lemm, Julie A.
; APPLICANT: O'Boyle, Donald R.
; APPLICANT: Nower, Peter
; TITLE OF INVENTION: HEPATITIS C VIRUS ASSAYS
; FILE REFERENCE: 10283 NP
; CURRENT APPLICATION NUMBER: US/11/119,330
; CURRENT FILING DATE: 2005-04-29
; PRIOR APPLICATION NUMBER: 60/567,270
; PRIOR FILING DATE: 2004-04-30

; PRIOR APPLICATION NUMBER: 60/568,590
; PRIOR FILING DATE: 2004-05-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 7989
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Replicon
US-11-119-330-1

Query Match 55.4%; Score 5318.4; DB 7; Length 7989;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 5654; Conservative 0; Mismatches 526; Indels 10; Gaps 1;

QY 3416 CCTTGGCCCATCACGGCTTACTCCCAACAAACGGGGGGGTACTTGGTTCATCATCAC 3475
DB 1800 CATGGCCGCTATTACGGCCTTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCAC 1859
QY 3476 TAGCTTCACAGGCGGGACAAAGAACACAGGTTCGAAGGGGAGGTTCAAGTGGTTTCTACCGC 3535
DB 1860 TAGCTTCACAGGCGGGACAAAGAACACAGGTTCGAAGGGGAGGTTCAAGTGGTTCTCCACGC 1919
QY 3536 AACAAATCTTTCTGTGCGACCTGCATCAACGGCGTGTGTGGACTGTCTACCATGGCGC 3595
DB 1920 AACAAATCTTTCTGTGCGACCTGCATCAATGGCGTGTGTGGACTGTCTATCATGGTGC 1979
QY 3596 TGGCTCGAAGACCTTACCGGTCCAAAAGGTCCAAATCACCCAAATGTACACCAATGTAGA 3655
DB 1980 CGGCTCAAAGACCTTTCGCGCCCAAGGGGCCAATCACCCAAATGTACACCAATGTGA 2039
QY 3656 CTTGCACTCGTTCGCGTGGCAGCGCCCGCGGGCGCGTCCATGACACCATCAGCTG 3715
DB 2040 CCAGGACTCGTTCGCGTGGCAAGCGCCCGCGGGCGGTTCTTGGACCATCGACCTG 2099
QY 3716 TGGCAGCTCGGACCTTTACTTTGGTCACGAGACATGCTGATGTCAATCCGGTGCGCCGCG 3775
DB 2100 CGGCAGCTCGGACCTTTACTTTGGTCACGAGGACATGCCGATGTCATTCGGTGGCGCG 2159
QY 3776 AGGCGACAGAGGGGAAGTCTACTCTCCCGAGCGCGTCTCTTACCTGAAAGGCTCCTC 3835
DB 2160 GGGCGACAGAGGGGAGCCTTCTCTCCCGAGCGCGTCTCTTACTTGAAGGGCTCTTC 2219
QY 3836 GGGTGGTCCATTGCTTTGCCCTTCGGGCGACGTCGTCGGGCTCTTCGGGCTGCTGTGTG 3895
DB 2220 GGGCGGTCCACTGCTCTGCCCTTCGGGCGACGCTGTGGGCACTCTTCGGGCTGCGGTG 2279
QY 3896 CACCGGGGGTTCGGAAGCGGTGGACTTTTCATACCCGTTGAGTCTATGGAACCTACCAT 3955
DB 2280 CACCGAGGGGTTCGGAAGCGGTGGACTTTGTATACCGTGCAGTCTATGGAACCTACAT 2339
QY 3956 GCGGTCTCGGTCTTTCACAGACAACTCAAACCCCGCGGTGTACCGGAGACATTCCAAGT 4015
DB 2340 GCGGTCCCGGTCTTTCACGGACAACTCGTCCCTTCGGCGGTACCGGACATTCACAGT 2399
QY 4016 GGCACATCTGCACGCTCCTACTGCGAGCGGACAGACCAAGTTCGGGCTGGGTATGC 4075
DB 2400 GGGCCATCTACACGCCCTCTCTGTAGCGGCAAGACACTAAGGTGCGCGCTGGGTATGC 2459
QY 4076 AGCCCAAGGGTACAAAGGTGCTCGTCCCTGAAACCCCGTTCGTCGCCCACTTTAGGGTTTGG 4135
DB 2460 AGCCCAAGGGTATAAGGTGCTTGTCTGAAACCCGTCGTCGCCCACTTTAGGGTTTGG 2519
QY 4136 GCGGTATATGTCCAAAGGACACCGGTATGACCCCTAAACATGAGAACTGGGGTAAAGACCAT 4195
DB 2520 GCGGTATATGTCTAAGGCACATGTTGTGACCCCTAAACATGAGAACTGGGGTAAAGACCAT 2579
QY 4196 TACCACGGCGGCTCCATTACGTACTCCACCTATGGCAAGTTCCTTGCAGCGGTGGCTG 4255
DB 2580 CACCACGGGTGCGCCCATCAGTACTCCACCTATGGCAAGTTCCTTGCAGCGGTGGCTG 2639
QY 4256 TTCTGGGGCGCCTATGACATCATATAATGTGAGTGGCCACTCACTGACTCGACTAC 4315

Db	2640	 CTCTGGGGCGCTATGACATCATATAATGTGATGAGTGCCACTCACTGACTCGACCAC	2699
Qy	4316	 CATCTTTGGGCATCGGCACACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGGCTCGTGGT	4375
Db	2700	 TATCTCTGGGCATCGGCACACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGGACTCGTGGT	2759
Qy	4376	 GCTCGCCACCGCTACACCTCCGGGATCGGTTACCGTGCACACCCCAATATCGAGAAAT	4435
Db	2760	 GCTCGCCACCGCTACCGCTCCGGGATCGGTCACCGTGCACATCCAAACATCGAGGAGT	2819
Qy	4436	 AGGCTGTCCAAACAATGGAGAGATCCCTCTTATGGCAAGGCATCCCATTTGAGGCCAT	4495
Db	2820	 GGCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGGCATCCCATCGAGACCAT	2879
Qy	4496	 CAAGGGGGGAGGCATCTCATTTTCTGCCATTTCCAAAGAAATGTGACGAGTCCGCCG	4555
Db	2880	 CAAGGGGGGAGGCATCTCATTTTCTGCCATTTCCAAAGAAATGTGATGAGCTGCCGC	2939
Qy	4556	 AAAGCTGACAGGCTTCGGACTGAACGCTGAGCATATTACCGGGGCTTTGATGTGCTCGT	4615
Db	2940	 GAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTTGATGTATCCGT	2999
Qy	4616	 CATACCGCCTATCGGAGACGTGCTGTGTGGCAACAGACGCTCTTAATGACGGGTTTCAC	4675
Db	3000	 CATACCAACTAGCGGAGACGTCTATTTGCTAGCAACGGACGCTCTTAATGACGGGCTTTAC	3059
Qy	4676	 CGGCGATTTTGACTCAGTGATCGACTGCATACATGTGTCAACCCAGACAGTGCAGTTTCA	4735
Db	3060	 CGGCGATTTTGACTCAGTGATCGACTGCATACATGTGTCAACCCAGACAGTGCAGTTTCA	3119
Qy	4736	 CTTGGATCCCACTTTACCAATTGAGACGACGACCGTGCCTCCCAAGACGCGTGTGCGGCTC	4795
Db	3120	 CTTGGAACCGNACTTTACCAATTGAGACGACGACCGTGCCTCCCAAGACGCGTGTGCGGCTC	3179
Qy	4796	 GCAACGCGAGGTTAGAACTGGCAGGGTTAGAGTGGCATCTACAGTTTGTGATCTCCAGG	4855
Db	3180	 GCAGCGGCAGGACGAGTCTGGTAGGGCAGGATGGGCAITTTACAGGTTTGTGACTCCAGG	3239
Qy	4856	 AGAACGGCCCTCGGACATGTTTCGATTTCTCGGTCTGTGTGAGTGTATGACCGGAGCTG	4915
Db	3240	 AGAACGGCCCTCGGACATGTTTCGATTTCTCGGTCTGTGTGAGTGTATGACCGGAGCTG	3299
Qy	4916	 TGCTTTGTTAGTCACTCAGCCCGCTGAGACCTCGGTTAGGTTGCGGCTTTACCTAAATAC	4975
Db	3300	 TGCTTTGTTAGTCACTCAGCCCGCGAGACCTCAGTTTGGTGGGCTTTACCTTAACAC	3359
Qy	4976	 ACCAAGGTTGCCGTCTTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTTTTCACAGCCCT	5035
Db	3360	 ACCAAGGTTGCCGTCTTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTTTTCACAGCCCT	3419
Qy	5036	 CACCCACATAGATGCCACTTCTGTGCCAGACTTAAACAGGCAGGAGACAACCTTTCCTTA	5095
Db	3420	 CACCCACATAGACGCCACTTCTTGTGCCAGACTTAAAGCAGGAGGAGACAACCTTCCCTTA	3479
Qy	5096	 CCTGTGTGCATATCAAGCTACAGTGTGGCCAGGGCTCAAGCTCCACTCCATCGTGGGA	5155
Db	3480	 CCTGTGTAGCATACCAAGGCTACGGTGTGGCCAGGGCTCAGGCTTCACTCCATCGTGGGA	3539
Qy	5156	 CCAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGTCAGCGGGCCAAACCCCTGCT	5215
Db	3540	 CCAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGTCAGCGGGCCAAACCCCTGCT	3599
Qy	5216	 GTATAGGCTAGAGCCGCTCAAAATAGGTCATCTCTACACACCCCATTAACATAACAT	5275
Db	3600	 GTATAGGCTGGAGCCGCTTCAAAACAGGTTTACTACCAACACCCCAATAACCAAAATACAT	3659
Qy	5276	 CATGCGATGATGTCGGCTGACCTGGAGGTGTACAGAGCACTTGGTGTCTGGTAGCGGG	5335
Db	3660	 CATGCGATGATGTCGGCTGACCTGGAGGTGTACAGAGCACTTGGTGTCTGGTAGCGGG	3719
Qy	5336	 AGTCTCTGAGCTTTGGCGGCATATGCTGCTGACGACAGGCAAGTGTGTCATTTGGGCGAG	5395

3720	AGTCTAGCAGCTCTGGCCGCGTATTGGCTCTGACAAACAGCAGCGTGGTCAATTTGTGGGCAG	3770
Qy	GATCATCTTTGTCGGGAAGCCAGAGCTGCTTCCCGACAGGGAAGTCTCTACCAAGAGTT	5455
Db		
3780	GATCATCTTGTCCGGAAGCGCGCCATCATTTCCGACAGGAAGTCTTTTACCGGAGTT	3839
Qy	CGATGAGATGGAAGAGTGTCCCTCACAACTTCTTACATCGACGAGGGAATCGAGCTCGC	5515
Db		
3840	CGATGAGATGGAAGAGTGGCCCTCACACCTCCCTTACATCGAACAGGGAATCGAGCTCGC	3899
Qy	CGAGCAATTCAAGCAAAAGCGCTCGGGTGTGTGCAAAAGGCCACCAAGCAAGCGGAGGC	5575
Db		
3900	CGAACAAATTCAAACAGAGGCATCGGTTGCTGCAACAGCCACCAGCAAGCGGAGGC	3959
Qy	TGCTGCTCCCGTGGAGTCCAAGTGGGAGCCCTTGATGACCTTCTGGGGCAGACACAT	5635
Db		
3960	TGCTGCTCCCGTGGAAATCCAAGTGGGGACCTTCGAAGCCCTTCTGGGCGAAGCATAT	4019
Qy	GTGGAAATTTTCATCAGCGGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAACCC	5695
Db		
4020	GTGGAAATTTTCATCAGCGGGATACAAATATTATAGCAGGCTTGTCCACTCTGCTGGCAACC	4079
Qy	CGCGATAGCATCATTGATGSCATTTACAGCTTCTATCACTAGCCCGCTCACCAACCAAAA	5755
Db		
4080	CGCGATAGCATCACTGATGSCATTCACAGCTCTATCAACAGCCCGCTCACCAACCAACA	4139
Qy	CACCTCCTGTTAAACATCTTGGGGGATGGGTGGTGCCAACTCGCTCCTCCAGCGC	5815
Db		
4140	TACCTCCTGTTAAACATCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGCGC	4199
Qy	TGCGTAGCCTTTGTTGGGCGCGGCATCGCGGAGCGGCTGTGTGGCAGCATAGCCCTTG	5875
Db		
4200	TGCTTCTGCTTTGATAGCGCCGCACTCGCTGAGCGGCTGTGTGGCAGCATAGCCCTTG	4259
Qy	GAAGTGCTCTGTGACATCTTGGCGGCTATGGGCGAGGGGTAGCCGGCGCACTCGTGGC	5935
Db		
4260	GAAGTGCTCTGTGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGCGCTCGTGGC	4319
Qy	CTTTAAGTGCATGAGCGGAGTGCCTCCACCGAGGACTCTGGTCAACTCTCCCTGC	5995
Db		
4320	CTTTAAGTGCATGAGCGGAGATGCCCTCCACCGAGGACTCTGGTTAACTCTCCCTGC	4379
Qy	CATCCTCTCTCGTGCCTTGTCTGCTCGGGTCTGTGCGCAGCAATCTGGTGGCA	6055
Db		
4380	TATCCTCTCCCTTGCGCCCTAGTCTGTGCGGGTCTGTGCGCAGCGATCTGGTGGCA	4439
Qy	CGTGGCGCCGGAGAGGGGCTGTGACGTGGATGAACCGGCTGATACGCTTCGCTCGCG	6115
Db		
4440	CGTGGCGCCAGGGAGGGGCTGTGACGTGGATGGAACCGGCTGATACGCTTCGCTCGCG	4499
Qy	GGGTAAACCACTCTCCCTACGCACTATGCTCTGAGAGCAAGCTCGACGAGTGTAC	6175
Db		
4500	GGGTAAACCACTCTCCCAACGCACTATGCTCTGAGAGCAAGCTCGACGAGTGTAC	4559
Qy	TCAGATCTCTTAGCTTACCATCACTCAACTGCTGAGCGGCTCCACAGTGGATTAA	6235
Db		
4560	TCAGATCTCTTAGCTTACCATCACTCAACTGCTGAGAGGCTTCAACAGTGGATTAA	4619
Qy	TGAGGACTGCTTAGCCCATGCTCCGGTCTGTGCTAAGGATGTTTGGGATTTGGATATG	6295
Db		
4620	CGAGACTGCTCCAGCCCATGCTCCGGTCTGTGCTAAGATGTTTGGGATTTGGATATG	4679
Qy	CACGGTGTGACTGACTTCAAGACTCTGGCTCCAGTCCAACTCTCTCGCGGGTTACCGGG	6355
Db		
4680	CACGGTGTGACTGATTTCAGACCTTGCTCCAGCTCCAAAGCTCTGCGCGGATTTGCCGGG	4739
Qy	AGTCCCTTCTTGTCATGCCAAACGGGGTACAAGGGAGTCTGGCGGGGGAGCGCATCAT	6415
Db		
4740	AGTCCCTTCTTCTCATGTCAACGTGGGTACAAGGGAGTCTTGGCGGGGCGAGCGCATCAT	4799
Qy	GCAACCACTCTGCCATCGGAGCAGCATCGCGGACATGTCAAAACCGTTCCATGAG	6475
Db		
4800	GCAACCACTCTGCCATGTGAGACAGATCACCGGACATGTGAATAACCGTTCCATGAG	4859

Db 7020 CTCTGCCCCCTGGGACCGGCCAAACAGAAATAGACTTGGAGTTGATACATCATG 7079
Qy 8696 TTCTCTCAATGTGTAGTCGCGCAGCATGTCATGGCAAAGGGTATACCTACCCCG 8755
Db 7080 CTCTCTCAATGTGTAGTCGCGCAGCATGTCATGGCAAAGGGTGTACTATCTCACCCG 7139
Qy 8756 TGACCCACACACCCCTTGCACGGCTGGTGGGAGACGTAGACACACTCCAATCAA 8815
Db 7140 TGACCCACACACCCCTTGCAGGGCTGGTGGGAGACGTAGACACACTCCAATCAA 7199
Qy 8816 CTCTTGCTAGGCAATATCATGATATGCGCCACCTATGGCAAGGATGATCTGAT 8875
Db 7200 TTCTTGCTAGGCAATATCATGATATGCGCCACCTTGTGGCAAGGATGATCTGAT 7259
Qy 8876 GACTCACTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTGAAGAGCCCTGGATGTCA 8935
Db 7260 GACTCATTTCTCCATCCTTCTAGCTCAGGAACAACCTTGAAGAGCCCTAGATGTCA 7319
Qy 8936 GATCTACGGGGCTGTCTCAATTTGAGCCACTTGCACCTACCTACCTCAGATCATTTGAAGACT 8995
Db 7320 GATCTACGGGGCTGTCTCAATTTGAGCCACTTGCACCTACCTACCTCAGATCATTTGAAGACT 7379
Qy 8996 CCATGGCTTAGCGCAATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGC 9055
Db 7380 CCATGGCTTAGCGCAATTTACACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGC 7439
Qy 9056 TTCAATGCTCAGGAACCTTGGGGTACCACCTTTCGGAACCTTGGAGACATCGGGCCAGAAG 9115
Db 7440 TTCAATGCTCAGGAACCTTGGGGTACCACCTTTCGGAACCTTGGAGACATCGGGCCAGAAG 7499
Qy 9116 TGTCCGGCTTAAGCTACTGTCCAGGGGGAGGGCCGCACTTGTGGAGATACCTCTTT 9175
Db 7500 TGTCCGGCTTAAGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTT 7559
Qy 9176 TAATGGGCAAGTAAAGCAAGCTTAACTCACTCCCAATCCCGCCCGCTCCAGCTGGA 9235
Db 7560 CAATGGGCAAGTAAAGCAAGCTTAACTCACTCCCAATCCCGCCCGCTCCAGCTGGA 7619
Qy 9236 TTGTCTGGCTGGTTTCGTGCTGGTTACAGCGGGGAGACATATATACAGCTGTCTCG 9295
Db 7620 TTATCCAGCTGGTTTCGTGCTGGTTACAGCGGGGAGACATATATACAGCTGTCTCG 7679
Qy 9296 TGCCCGACCCCGCTGGTTTCGTGCTGGTTACAGCTACTCTACTTTCTGTAGGGTAGGCATTA 9355
Db 7680 TGCCCGACCCCGCTGGTTTCGTGCTGGTTACAGCTACTCTACTTTCTGTAGGGTAGGCATTA 7739
Qy 9356 CTGTCTCCCAACCGATGAACGGGAGCTAACCACTCCAGGCTTAAGCC----- 9405
Db 7740 TCTACTCCCAACCGATGAACGGGAGCTAACCACTCCAGGCTTAAGCC----- 7799
Qy 9406 ATTTCTCTGTTTCTTCT 9465
Db 7800 TTTTCCCTTCTCTTT 7859
Qy 9466 TCTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9525
Db 7860 TTTTCTCTTCTCTTT 7919
Qy 9526 CTAGCTGGAAGGTTCGTGAGCGCGATGACTGACAGAGAGTGTGATCTGGCCCTCTCTG 9585
Db 7920 CTAGCTGGAAGGTTCGTGAGCGCGTGTGACTGACAGAGAGTGTGATCTGGCCCTCTCTG 7979
Qy 9586 CAGATCATGT 9595
Db 7980 CAGATCAAGT 7989

RESULT 9
US-11-111-686-1
; Sequence 1, Application US/1111686
; Publication No. US20050260221A1
; GENERAL INFORMATION:

; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPI
; FILE REFERENCE: 0342/1H395US1
; CURRENT APPLICATION NUMBER: US/11/111,686
; CURRENT FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: US/10/005,469
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV replicon I377/NS3-3'UTR
US-11-111-686-1

Query Match 55.4%; Score 5318.4; DB 7; Length 7992;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 5654; Conservative 0; Mismatches 526; Indels 10; Gaps 1;
Qy 3416 CCTTGGCCCATCACGGCTTACTCCAAACAAACGCGGGCGTACTTGGTTGCATCATCAC 3475
Db 1800 CATGGCGCTATTACGGCTTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCAC 1859
Qy 3476 TAGCCTCACAGCGCGGACAAAGAACAGAGTCAAGGGAGGTTCAAGTGGTTTCTACCGC 3535
Db 1860 TAGCCTCACAGCGCGGACAAAGAACAGAGTCAAGGGAGGTTCAAGTGGTCTCCACCGC 1919
Qy 3536 AACAAATCTTCTGGGACCTTGATCAACAGCGGGTGTCTGACGTCTACCATGCGC 3595
Db 1920 AACAAATCTTCTGGGACCTTGATCAACAGCGGGTGTCTGACGTCTATCATGTTGC 1979
Qy 3596 TGCTCGAAGACCTTAGCGGTCCAAAGGTCCAAATCACCACCAATGTACACCAATGTAGA 3655
Db 1980 CGGTCAAGACCTTTCGCGGCCCAAGGGCCCAATCACCACCAATGTACCAATGTAGA 2039
Qy 3656 CTGGACCTCTGGTGGCAGCGGCCCGCGGGCGCTCCATGACACCATGACAGCTG 3715
Db 2040 CCAGGACCTCTGGTGGCAGCGGCCCGCGGGCGGTTCCTTTGACACCATGACAGCTG 2099
Qy 3716 TGGCAGCTCGGACTTTTACTTTGTCAGAGACATGTGTATGTCTATTCGGTGGCCGCG 3775
Db 2100 CGGCGCTCGGACTTTTACTTTGTCAGAGGCTTGCAGTGTCAATTCGGTGGCCGCG 2159
Qy 3776 AGGCGACAGCGGGAGTCTACTCTCCCGCAGCGCGCTCTCTACCTGAAAGGCTCTC 3835
Db 2160 GGGCGACAGCGGGGAGCTTCTCTCCCGCAGCGCGCTCTCTTCTTGAAGGCTCTTC 2219
Qy 3836 GGGTGGTCCATTTGCTTGGCCCTTGGGGCAGCTGTGGGGGTCTTCCGGGCTGCTGTG 3895
Db 2220 GGGGGTCTCACTGCTCTGCCCCCTCGGGCAGCTGTGGGCACTTTTCCGGCTGCGGTG 2279
Qy 3896 CACCGGGGGGTTCGCAAGCGGTGGACTTCAATACCGTTGAGTCTATGGAACACTACCAT 3955
Db 2280 CACCGGAGGGTTTGCAGAGCGGTGGACTTTGTACCCGCTCGAGTCTATGGAACACTAT 2339
Qy 3956 GCGGTCTCCGGTCTTTCACAGACAACTCAACCCCCCGGCTGTACCGGAGACATTCAGT 4015
Db 2340 GCGGTCTCCGGTCTTTCACGACAACTCGTCCCCCTCCCGCGGTACCGGAGACATTCAGT 2399
Qy 4016 GGCACATCTGCAGCTCTTACTGCGCGGCAAGAGACCAAAAGTGGCGGCTGCGTATGC 4075
Db 2400 GGCCCATCTACACCCCTTACTGTAGGGCAAGAGCACTAAGTGGCGGTGGTATGC 2459
Qy 4076 AGCCCAAGGGTACAAAGGTGTCTGCTTGAACCCCGTTCGGTCCGCCACCTTAGGGTTGG 4135
Db 2460 AGCCCAAGGGTATAAAGGTGTCTGCTTGAACCCCGTTCGGTCCGCCACCTTAGGGTTGG 2519
Qy 4136 GGCCTATATGTCCAAAGGACACCGGTATCGACCCCTAAACATCAAGAACTGGGGTAAGGACCAT 4195

Db 2520 |||||GGCGTATATGTCTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGGTAAGACCAT 2579
QY 4196 TACACAGGGCGGCTCCATTACGTACTCCACCTATCGCAAGTTCCTTGCAGCGGTGGCTG 4255
Db 2580 CACCAACGGGTGCCCCATCACTACGTACTCCACCTATCGCAAGTTCCTTGCAGCGGTGGTTG 2639
QY 4256 TTCTGGGGCGGCTATGACATCAATAATGTGATGAGTGCACCTCAACTGACTCGACTAC 4315
Db 2640 CTCTGGGGCGGCTATGACATCAATAATGTGATGAGTGCACCTCAACTGACTCGACCCAC 2699
QY 4316 CATCTTGGGCATCGGCACAGTCTCTGGACCAAGCGAGACGGTGGAGCGCGCTCGTCTG 4375
Db 2700 TATCTGGGCATCGGCACAGTCTCTGGACCAAGCGAGACGGTGGAGCGCGACTCGTCTG 2759
QY 4376 GCTCGGCACCGGTACACCTCGGGATCGGTGTACCGTGCCACACCCCAATATCGAGGAAT 4435
Db 2760 GCTCGGCACCGGTACCGCTCGGGATCGGTGTACCGTGCCACACCCCAATATCGAGGAGT 2819
QY 4436 AGGCTGTCCAAATGGAGAGATCCCTTCTATGCAAGGACATCCCATTTAGGCGCAT 4495
Db 2820 GGCTCTGTCCAGACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATTCGAGACCAT 2879
QY 4496 CAAAGGGGGAGGACATCTCATTTTCTGCCATTTCCAAAGAGAAATGTGACGACTCGCGC 4555
Db 2880 CAAAGGGGGAGGACCTCATTTTCTGCCATTTCCAAAGAGAAATGTGATGAGTCTCGCGC 2939
QY 4556 AAAGCTGACAGCGCTCGGACTGAAAGCTGTGACATATTAACGGGGCTTGTATGTCTCGT 4615
Db 2940 GAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATGTATCGT 2999
QY 4616 CATACCGCTATCGGAGAGCTGTGTGCTGCGCAACAGACGCTCTAATGACGGGTTTCAC 4675
Db 3000 CATACCACTAGCGGAGAGCTCATTTGCTGAGCAACGAGAGCTCTAATGACGGGCTTTAC 3059
QY 4676 CGGCGATTTTGAATCAGTGATCGACTGCAATACATGTGTACCCAGACAGTGCAGTTCCAG 4735
Db 3060 CGGCGATTTTGAATCAGTGATCGACTGCAATACATGTGTACCCAGACAGTGCAGTTCCAG 3119
QY 4736 CTTTGGATCCACCTTACCATTTAGAGACGACGCGTGCCCAAGACGCGGTGTGCGGCTC 4795
Db 3120 CTTGACCCGACCTTACCATTGAGACGACGCGTGCCCAAGACGCGGTGTGACGCTC 3179
QY 4796 GCAACGGCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGTTTGTACATCCAGG 4855
Db 3180 GCAGCGCGGAGGACGACTGGTAGGGGAGGATGGGCATTTACAGTTTGTGACTCCAGG 3239
QY 4856 AGAACGGCCCTCGGGCATGTTGATTTCTTCGGTCTCTGTGTGAGTGTATGACGCGGGCTG 4915
Db 3240 AGAACGGCCCTCGGGCATGTTGATTTCTTCGGTCTCTGTGTGAGTGTATGACGCGGGCTG 3299
QY 4916 TGGTTGGTATGAGCTCACGCGGCTGAGACCTCGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAG 4975
Db 3300 TGGTTGGTATGAGCTCACGCGGCTGAGACCTCAGTTAGGTTAGGTTAGGTTAGGTTAGGTTAG 3359
QY 4976 ACCAGGTTGCGGCTGCGAGACCATCTGAGTTCTGGGAGAGGCTCTTACAGGCT 5035
Db 3360 ACCAGGTTGCGGCTGCGAGACCATCTGAGTTCTGGGAGAGGCTCTTACAGGCT 3419
QY 5036 CACCCACATAGATGCCACTTCTGTGCCAGACTTAAACAGGAGGAGACAACTTTCCTTAA 5095
Db 3420 CACCCACATAGATGCCACTTCTGTGCCAGACTTAAACAGGAGGAGACAACTTTCCTTAA 3479
QY 5096 CTTGTTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGTGTCAACCTTCAATCGTGGGA 5155
Db 3480 CTTGTTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGTGTCAACCTTCAATCGTGGGA 3539
QY 5156 CCAATGTGGAGTGTCTCATACGGCTGAAACCTACACCTGCAAGGGCCAAACCTTCTGCT 5215
Db 3540 CCAATGTGGAGTGTCTCATACGGCTTAAAGCCCTACGCTGCAAGGGCCAAACCTTCTGCT 3599
QY 5216 GTATAGGCTAGGAGCGGTCCTCAAAATGAGGTCTATCTTCAACACACCCCATTAACATACAT 5275

Db 3600 GTATAGGCTGGAGCCGTTCAAAACGAGTTTACTACACACACCCCCATAACAAATACAT 3659
QY 5276 CATGGCATGATGTGCGTACCTGGAGGTGCTCACTAGCACCTGGGTGCTGTAGGCGG 5335
Db 3660 CATGGCATGATGTGCGTACCTGGAGTGTGCTCAAGACACCTGGGTGCTGTAGGCGG 3719
QY 5336 AGTCTTTCAGCTTTTGGCCGCATCTCCCTGACACAGGCAAGTGTGTTTATTTGGGCGAG 5395
Db 3720 AGTCTTAGCAGTCTGGCCGCTATTGCTGACACAGGCAAGTGTGTTTATTTGGGCGAG 3779
QY 5396 GATCATCTTGTTCGGGAAGCAGCTGTCTTCCGACAGGGAAGTCTCTTACAGGAGTT 5455
Db 3780 GATCATCTTGTTCGGGAAGCAGCTGTCTTCCGACAGGGAAGTCTCTTACAGGAGTT 3839
QY 5456 CGATGAGTGAAGAGTGTGCTCACAACCTTCTTACATCGACAGGGAAGTGTGAGTCTGCG 5515
Db 3840 CGATGAGTGAAGAGTGTGCTCACAACCTTCTTACATCGACAGGGAAGTGTGAGTCTGCG 3899
QY 5516 CGAGCAATTCAGCAAAAGGCGCTCGGGTTTGTGCAACCGCCACCAAGCAAGCGAGGCG 5575
Db 3900 CGACAAATTCAGCAAAAGGCAATCGGGTTGCTGCAACAGCCACCAAGCAAGCGAGGCG 3959
QY 5576 TGCTGCTCCGCTGTGTGAGTCCAAAGTGTGAGGCGCTTGTGAGACCTTCTGGGCGAAGCA 5635
Db 3960 TGCTGCTCCGCTGTGTGAGTCCAAAGTGTGAGGCGCTTGTGAGACCTTCTGGGCGAAGCA 4019
QY 5636 GTGGAATTCATACGCGGAATACAGTACCTAGTACGAGGCTTATCCACTCTGCTGGAACCC 5695
Db 4020 GTGGAATTCATACGCGGAATACAGTATTTAGAGGCTTGTCCACTCTGCTTGGCAACCC 4079
QY 5696 CGCGATAGCATCATTTGATGAGTATTAACAGCTTCTATCACTAGCCGCTCACCACCACAAA 5755
Db 4080 CGCGATAGCATCATTTGATGAGTATTAACAGCTTCTATCACTAGCCGCTCACCACCACAAA 4139
QY 5756 CACCTCTCTGTTTAACTTTTGGGGGAGTGGTGTGCTGCCCAACTCTCTCTCCAGGCG 5815
Db 4140 TACCTCTCTGTTTAACTTTTGGGGGAGTGGTGTGCTGCCCAACTCTCTCTCCAGGCG 4199
QY 5816 TGGCTCAGCTTTCTGTTGGGCGCGCATCGCGAGGCTGTGCGACATAGGCTTGG 5875
Db 4200 TGGCTCAGCTTTCTGTTGGGCGCGCATCGCTGGAGGCTGTGCGACATAGGCTTGG 4259
QY 5876 GAAAGTGTCTGTGACATCTTTGGCGGCTATGGGGAGGGGTAGCGCGGCACTCTGTGCG 5935
Db 4260 GAAAGTGTCTGTGACATCTTTGGCGGCTATGGAGAGGGGTAGCGCGGCACTCTGTGCG 4319
QY 5936 CTTTAAAGTGTATGACGCGGAGTGTGCTTCCACGAGGACCTGTGTTAACTTCTCTCTGCG 5995
Db 4320 CTTTAAAGTGTATGACGCGGAGATGCTTCCACGAGGACCTGTGTTAACTTCTCTCTGCG 4379
QY 5996 CATCTCTCTCTGTTGCGCTGTGCTGCGGCTGTGCGAGCAATACTGCGTCCGCA 6055
Db 4380 TATCTCTCTCTGTTGCGCTGTGCTGCGGCTGTGCGAGGATACTGCGTCCGCA 4439
QY 6056 CGTGGGCGCGGAGAGGGGCTGTGCTGAGTGAACCGGCTGTAGGCTTCTGCTTCCGCG 6115
Db 4440 CGTGGGCGCGGAGAGGGGCTGTGCTGAGTGAACCGGCTGTAGGCTTCTGCTTCCGCG 4499
QY 6116 GGGTAAACAGCTCTCTTCAAGCTATGTGCTGAGAGGAGCTGTGAGGATTTGGGATTTGATG 6175
Db 4500 GGGTAAACAGCTCTCTTCAAGCTATGTGCTGAGAGGAGCTGTGAGGATTTGGGATTTGATG 4559
QY 6176 TCAGATCTCTCTAGCCTTACATCACTGTGAGGCTTCCAGGCTTCCAGGATTTAA 6235
Db 4560 TCAGATCTCTCTAGCCTTACATCACTGTGAGGCTTCCAGGATTTAA 4619
QY 6236 TGAGGACTGTCTAGCCTTACATCACTGTGAGGCTTCCAGGATTTAA 6295
Db 4620 TGAGGACTGTCTAGCCTTACATCACTGTGAGGCTTCCAGGATTTAA 4679
QY 6296 CACGCTGTGATGATCTTCAAGACCTGTGCTCAGTCTCAAACTCTCTGCGCGGTTTACCGGG 6355
Db 4680 CACGCTGTGATGATCTTCAAGACCTGTGCTCAGTCTCAAACTCTCTGCGCGGTTTACCGGG 4739

Qy	6356	AGTCCCTTTCTGTCTATGCCAACGCGGGTACAAGGAGTCTGGCGGGGAGCGCATCAT	6415
Db	4740		
Qy	6416	GCAAAACACCTGCGCCATCGGAGACACAGATCGCGGACATGTCAAAAAACGGTTCATGAG	6475
Db	4800	GCAAAACACCTGCGCCATGTGGAGACAGATCAACGACATGTGAAAAACGGTTCATGAG	4859
Qy	6476	GATCGTAGGCGCTAGAACCTGACACACAGTGGCACGGAAAGTTCCTCCCATCAACGCATA	6535
Db	4860	GATCGTGGGCGCTAGGACCTGTATACACGTGGCATGGAAACATTCCTCATTAACCGTGA	4919
Qy	6536	CACCAACGGACCTTGACACACCTCCCGCGCGCCAACTATTTCAGGCGCGTATGGCGGT	6595
Db	4920	CACCAACGGCGCCCTGCACGCGCTCCCGCGCGCCAAATTTATTCTAGGCGGTGTGGCGGT	4979
Qy	6596	GGCTGCTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTTCCACTACGTACGCGGCAAT	6655
Db	4980	GGCTGCTGAGGAGTACGTGGAGGTTACGCGGTGGGGATTTTCCACTAGCTGACGCGCAT	5039
Qy	6656	GACCACTGACAAACGTAAAGTGCACAGTTCGGGCCCCCGAATTTCTTACGAGGT	6715
Db	5040	GACCACTGACAAACGTAAAGTGCCCGTGTAGGTTTCGGGCCCCCGAATTTCTTACAGAGT	5099
Qy	6716	GGATGAGTTCGGTTCACAGGTACGCTCCGGCGTGCMAACCTCTTCTACGGGAGACGT	6775
Db	5100	GGATGCGGTTCGGTTCACAGGTACGCTCAGCGTGCMAACCTCTTCTACGGGAGAGGT	5159
Qy	6776	CACGTTCCAGTTCGGGCTCAACCAATACTTGGTGGGTGCGAGTCCCATGTGAGAGCCGA	6835
Db	5160	CACATTCCTGGTTCGGGCTCAATCAATACCTGGTGGGTACAGCTCCCATGCGAGCCGA	5219
Qy	6836	ACCGGACGTAAACAGTGCTTACTTCATGCTCAACGATCCCTCCCACTTACAGCAGAC	6895
Db	5220	ACCGGACGTAGCAGTGCTCACTTCCATGCTCACCGACCCCTCCCACTTACGGCGAGAC	5279
Qy	6896	GGCTAAGCGTAGGCTGAGGGTCTCCCGCTTTAGCCAGCTCATCAGCTAGCCCA	6955
Db	5280	GGCTAAGCGTAGGCTGCGCAGGGATCTCCCGCTCTTGGCCAGCTCATCAGTAGCCA	5339
Qy	6956	GTTCTCTCGCGCTTTTGAAGGCGACATGCATACCCACATGACTCCCGGACGCTGA	7015
Db	5340	GCTCTCTGCGCTTCTTTGAAGGCAACATGCACCTACCCGTCATGACTCCCGGACGCTGA	5399
Qy	7016	CCTCATCGAGGCCAACCTCTTGTGGCGGACAGAGATGGGCGGAAAATCATCATCTCGGTGGA	7075
Db	5400	CCTCATCGAGGCCAACCTCTGTGGCGGACAGAGATGGGCGGAAAATCATCACCGCGCTGA	5459
Qy	7076	GTACAGAGTAAGTAGTAAATCTTGGAATCTTTTGGACCTTTTCGACCGCTCACGCGAGGGGATGA	7135
Db	5460	GTACAGAAATAAGGTAGTAAATTTTGGACCTTTTCGACCGCTTCACGCGGAGGAGATGA	5519
Qy	7136	GAGGAGATATCCGTTCGCGGCGGAGATCCTCGGAAAATCCAGGAAGTTCCCTCAGCGTT	7195
Db	5520	GAGGAGATATCCGTTTCGCGGCGGAGATCCTCGGAGGTTCAGGAAATTCCTCGAGCGAT	5579
Qy	7196	GCCCATATGGGACGCCCGGACTCAATCTCTCACTGCTAGGTCCTTGGGAAGGACCGGGA	7255
Db	5580	GCCCATATGGGACGCCCGGATTAACACCTCCACTGTTAGTCTCTGGAAGGACCCGGA	5639
Qy	7256	CTACGTCCCTTCGGGTGTACAGGATGCCCATTCGCACTTACCAAGGCTCTCCATATACC	7315
Db	5640	CTACGTCCCTCTCAGTGTGTACAGGGTGTTCATATGCGCGCTGCCAAGGCCCTCCGATATACC	5699
Qy	7316	ACCTCCACGGAGAAAGGACGGTTGCTCTCACAGAAATCCAAATGTGTCTTCTGCTTGGC	7375
Db	5700	ACCTCCACGGAGAAAGGACGGTTGTCTGTACAGAAATCTACCGGTGTCTTCTGCTTGGC	5759
Qy	7376	GGAGCTTCGCCACTAAGACCTTTCGGTAGCTTCGGGATTCGTCGGCCGTGTGATAGCGCACGGC	7435
Db	5760	GGAGCTTCGCCACAAAGACCTTTCGGCAGCTTCGAAATGTCGGCCGTGCACAGCGCACGGC	5819

Qy	7436	GACGGCCTTCTCTGACCTGTGGCTTCCAGCAGCGTGACAAAGATCCGAGCTTGACGTCTGA	74395
Db	5820	AACGGGCTCTCTCTGACAGCGCTTCCAGCAGCGCGGATCCGAGCTTGAGTCTGAGTCTGA	5879
Qy	7496	CTCCTCCATGCGCCCTTGAAGGGGAGCGCGGGGACCCCGATCTCAGCGACGGGCTCTTG	7555
Db	5880	CTCTCATGCGCCCTTGAAGGGGAGCGCGGGGATCCGATCTCAGCGACGGGCTCTTG	5939
Qy	7556	GTCTACCGTGAGTGAGGAGGCTAGTGAGGATGTGCTGTGCTCCTCAATGTCTTATACGTG	7615
Db	5940	GTCTACCGTAAGCGAGGAGGCTAGTGAGGACGTGCTGTGCTGTGATGTCTTACACATG	5999
Qy	7616	GACAGGGCCCTGATCAGCGCATGCGCTCGGAGGAAAGTAAAGTGGCCCATCAACCCGTT	7675
Db	6000	GACAGGGCCCTGATCAGCGCATGCGCTCGGAGGAAACCAAGTGGCCCATCAATGCACT	6059
Qy	7676	GAGCAACTTTTGTGCTGCATCACACAACATGTGCTACGCCACAACATCCCGCAGCGCAAG	7735
Db	6060	GAGCAACTTTTGTGCTGCATCACACAACATTTGTTCTATGCTAAACATCTCGAGGCGCAAG	6119
Qy	7736	CCTCGGCGAAGAAGGTCACTTTGACAGATGCAAGTCTCGATGATCATTAACCGGGA	7795
Db	6120	CCTCGGCGAAGAGGTCACTTTGACAGATGCAAGTCTCGATGATCATTAACCGGGA	6179
Qy	7796	CGTACTCAAGGAGATGAAGCGAAGGCGTCCACAGTTTAAGGCTTAAGCTTCTATCTATAGA	7855
Db	6180	CGTCTCAAGGAGATGAAGCGAAGGCGTCCACAGTTTAAGGCTTAAGCTTCTATCTCTGGA	6239
Qy	7856	GGAGGCTCGAAGTGAAGCGGCGCCACATTCGGCCCAATCCAAATTTGCTATGGGCGAAA	7915
Db	6240	GGAAGCTGTAAAGTGAAGCGGCGCCACATTCGGCCCAATTCAAATTTGCTATGGGCGAAA	6299
Qy	7916	GGAGCTCCGGAACCTATCCAGCAGGCGCGTTTAACCATCCGCTCCGTTGGGGAGGACTT	7975
Db	6300	GGAGCTCCGGAACCTATCCAGCAGGCGCGTTTAACCATCCGCTCCGTTGGGGAGGACTT	6359
Qy	7976	GCTGGAAAGACCTGAAACACAAATTGACACCATCATGTCGCAAAAAGTGAGGTTTCTG	8035
Db	6360	GCTGGAAAGACCTGAGACACCAATTGACACCATCATGTCGCAAAAAGTGAGGTTTCTG	6419
Qy	8036	CGTCCAACGAGAGAGGAGCGCGCAGCTCGGCTTATCGTATTTCCAGACCTGGG	8095
Db	6420	CGTCCAACGAGAGAGGAGCGCGCAGCTCGGCTTATCGTATTTCCAGACCTGGG	6479
Qy	8096	AGTTCTGTGTATGCGAAGAAGATGGCGCTTTACGACGTGTCTCCACCTCTCCAGGCGGT	8155
Db	6480	GGTTCTGTGTGCGAAGAAATGGCGCTTTACGATGTGTCTCCACCTCTCCAGGCGGT	6539
Qy	8156	GATGGGCTCTCATACGGAATTCAAATACTCCCCAAGCAGCGGTTCGAGTTCTGTGTGAA	8215
Db	6540	GATGGGCTCTTCATACGGAATTCAAATACTCTCTGGGACAGCGGTTCGAGTTCTGTGTGAA	6599
Qy	8216	TACCTGGAATTCAAAGAAATGCCCCATTTGGGCTTCTCATATGACACCCGCTGTTTGACTC	8275
Db	6600	TGCTCTGAAAGCGAAGAAATGCCCTATATGGGCTTCGCATATGACACCCGCTGTTTGACTC	6655
Qy	8276	AACGCTCACTCAGAGTGACATTCGTTGTTGAGAGTCAATTTACCAATGTTGTGACTTGGC	8335
Db	6660	AACGCTCACTCAGAGTAATGATCCCGTTGTTGAGAGTCAATCTACCAATGTTGTGACTTGGC	6719
Qy	8336	CCCGAGGCCAGACAGGCCATAAGTCTGCTCACAGAGCGGCTTTACATCGGGGGTCCCT	8395
Db	6720	CCCGAGGCCAGACAGGCCATAAGTCTGCTCACAGAGCGGCTTTACATCGGGGGTCCCT	6779
Qy	8396	GACTAACTCAAAGGGCAGAACTCGGTTATTCGCGGTGCGCGCAAGTGGCGTCTGAC	8455
Db	6780	GACTAAATTTAAAGGGCAGAACTCGGCTATCGCGGTGCGCGCGAGCGGTGTACTGAC	683
Qy	8456	GACTAGCTGCGGTAAATACCTCTCAATGTTACTTGAAGGCACTGCAGGCTGTCCAGCTGC	8519
Db	6840	GACCACTGCGGTAAATACCTCTCAATGTTACTTGAAGGCGCTGCGGCTGTCCAGCTGC	689
Qy	8516	AAAGCTCCAGACTCAGCATGCTCGTGAAACGAGACGACCTTGTGTTATCTGTGAAAG	8579

Db 6900 ||||| GAAGCTCCAGGACTGCAGATCTCGTATGCGGAGACGACCTTGTCTATCTGTGAAG 6959
Qy 8576 CCGCGGAACCCAGGAGGATCGCGGCGCTTACGAGCTTTCACGGAGGCTATGACTAGTA 8635
Db 6960 CCGCGGGACCCAGAGGAGGAGCGAGCTACGGGCTTTCACGGAGGCTATGACTAGTA 7019
Qy 8636 TTCCGCCCCCGGGGATCGCCCCAACCAAGATACGACCTGGAGCTGATAACATCATG 8695
Db 7020 CTCTGCCCCCTGGGACCGCCCCAACCAAGATACGACCTGGAGGTTGATAACATCATG 7079
Qy 8696 TTCCTCCATGTGTAGTCGCCAGATGTCATCTGGGAAAAGGTTATACTACCTCACCG 8755
Db 7080 CTCCTCCAATGTGTAGTCGCCAGATGTCATCTGGGAAAAGGTTATACTCTCACCG 7139
Qy 8756 TGACCCCAACCCCTTTCAGCGGCTGCGTGGGAGACAGCTAGACACACTCCCAATCAA 8815
Db 7140 TGACCCCAACCCCTTTCAGCGGCTGCGTGGGAGACAGCTAGACACACTCCCAATCAA 7199
Qy 8816 CTCCTGGCTAGGCAATATCATCATGTATGCGGCCACCTTATGGCAAGGATGATTCTGAT 8875
Db 7200 TTCCTGGCTAGGCAATATCATCATGTATGCGGCCACCTTATGGCAAGGATGATTCTGAT 7259
Qy 8876 GACTCACTTTTCTCATCTCTTCTAGCTCAAGCAACTTGAAAAAGCCCTGGATTGCA 8935
Db 7260 GACTCATTTCTCTCATCTCTTCTAGCTCAGGAACAACCTTGAAAAAGCCCTAGATTGCA 7319
Qy 8936 GATCTACGGGCTTCTACTCCATTCAGCCACTTGACCTACCTCAGATCATTCAGGACT 8995
Db 7320 GATCTACGGGCTTCTACTCCATTCAGCCACTTGACCTACCTCAGATCATTCAGGACT 7379
Qy 8996 CCATGGCTTTAGGCAATTTACACTCCAGTTTACTCTCCAGGTGAGATCAATAGGCTGCG 9055
Db 7380 CCAATGGCTTTAGGCAATTTTCTACTCCATGTTACTCTCCAGGTGAGATCAATAGGCTGCG 7439
Qy 9056 TTTCATGCTCTAGGAACTTTGGGGTACCACCTTTCGGAACCTGGAGACATTCGGGCGAGAG 9115
Db 7440 TTTCATGCTCTAGGAACTTTGGGGTACCACCTTTCGGAACCTGGAGACATTCGGGCGAGAG 7499
Qy 9116 TGTCCGCGTACTGTCTCCAGGGGGAGGCGGCACTTGTGGCAGATACCTCTT 9175
Db 7500 TGTCCGCGTACTGTCTCCAGGGGGAGGCGGCACTTGTGGCAGATACCTCTT 7559
Qy 9176 TAACTGGGCACTAAGGACCAAGCTTAACTCACTCCAATCCGCGCGCTCCAGCTGA 9235
Db 7560 CAACTGGGCACTAAGGACCAAGCTCAAACCTCACTCCAATCCGCGCTCCAGTTGA 7619
Qy 9236 CTTGTCTGGCTGGTTCGTGCTGTTACAGCGGGGAGACATATATCAAGCCTGTCTCG 9295
Db 7620 TTTATCCAGCTGGTTCGTGCTGTTACAGCGGGGAGACATATATCAAGCCTGTCTCG 7679
Qy 9296 TGCCGACCCCGTGGTTTCGTTGCTACTCTCTACTTTCTGTAGGGTAGGCAATTA 9355
Db 7680 TGCCGACCCCGTGGTTTCATGTGGTGCCTACTCTCTACTTTCTGTAGGGTAGGCAATTA 7739
Qy 9356 CTGTCTCCCAACCAAGTAAAGGGAGCTAAACACTTCCAGGCTTAAAGCC----- 9405
Db 7740 TCTACTCCCCACCGATGAACGGGAGCTAAACACTTCCAGGCAATAGGCCATCTGTITT 7799
Qy 9406 ATTTCCTGTTT 9465
Db 7800 TTTTCCCTT 7859
Qy 9466 TCTTTTTTTCCTTTCTTTTCCGTTCCGTTCTTAATGGTGGCTCCATCTTAGCCCTAGTCAAG 9525
Db 7860 TTTTTTCTCTTTTTTCTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7919
Qy 9526 CTAGCTGTAAAGGTCGCTGAGCCGATGACTGACAGAGAGTGTGATAGTGCCTCTCTG 9585
Db 7920 CTAGCTGTAAAGGTCGCTGAGCCGTTGACTGACAGAGAGTGTGATAGTGCCTCTCTG 7979
Qy 9586 CAGATCATGT 9595
|||||

Db 7980 CAGATCAAGT 7989
RESULT 10
US-11-111-686-6
; Sequence 6, Application US/11111686
; Publication No. US20050260221A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPI
; FILE REFERENCE: 0342/IH395US1
; CURRENT APPLICATION NUMBER: US/11/111.686
; CURRENT FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: US/10/005.469
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245.866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Replicon from cell line HCVr24
US-11-111-686-6
Query Match 55.4%; Score 5316.8; DB 7; Length 7992;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 5653; Conservative 0; Mismatches 527; Indels 10; Gaps 1;
Qy 3416 CTTTGGCGCCATCAGCGCTACTCCCAACAAAGCGGGCGGTACTTGGTTGCATCATCAC 3475
Db 1800 CATGGCGCTATTACGGCTACTCCCAACAGACGAGCGCTACTTGGCTGCATCATCAC 1859
Qy 3476 TAGCCTCAAGCGCGGCAAGAACACAGGTCCGAGGGAGGTTCAAGTGGTTTACCGC 3535
Db 1860 TAGCCTCAAGCGCGGCAAGAACACAGGTCCGAGGGAGGTTCCAAGTGGTTCTCACCGC 1919
Qy 3536 AACACAATCTTCTGGCGACTTGCATCAAGCGGTGTCTGGACTGTCTACCATGGCGC 3595
Db 1920 AACACAATCTTCTGGCGACTTGCATCAATGGCGGTGTCTGGACTGTCTATCATGGTGC 1979
Qy 3596 TGGCTCGAAGACCTTAGCGGTCCAAAGGTCCTCAATCACCACCAATGTACCAATGTAGA 3655
Db 1980 CGGCTCAAAGACCTTTCGCGCCCAAGGCGCAATCACCACCAATGTACCAATGTAGA 2039
Qy 3656 CTTGGACCTCGTGGCTGGCAGCGCCCCCGGGGCGGCTCCATGACACCATGACGCTG 3715
Db 2040 CCAGGACCTCGTGGCTGGCAGCGCCCCCGGGGCGGCTCCATGACACCATGACGCTG 2099
Qy 3716 TGGGAGCTCGGACCTTTACTTGGTACGAGACATGCTGATGTCTATCCGGTGGCGGCG 3775
Db 2100 CGGACAGCTCGGACCTTTACTTGGTACGAGGATCCGATGTCTTCCGGTGGCGGCG 2159
Qy 3776 AGGCGACAGCGGGAAGTCTACTCTCCCGACGCGCGCTCTCTACTCTGAAAGGCTCTTC 3835
Db 2160 GGGCGACAGCGGGAAGGCGCTACTCTCCCGACGCGCGCTCTCTACTCTGAAAGGCTCTTC 2219
Qy 3836 GGGTGGTCCATTTGCTTTTCCCTTCGGGGCAGCGTCTGGGCGCTCTTCCGGGCTGTGTGTG 3895
Db 2220 GGGCGGTCCACTGTCTTCCCTTCGGGCGACGCTGTGGGCGCTCTTCCGGGCTGTGTGTG 2279
Qy 3896 CACCGGGGGGTTCGGAAGGCGGTGACTTCAATCCCGTTGAGTCTATGGAACTTACCAT 3955
Db 2280 CACCGGAGGGGTTCGGAAGGCGGTGACTTGTATCCCGCTCGAGTCTATGGAACTTACCAT 2339
Qy 3956 GCGGTCTCCGCTTTTCAGAGACCACTCAACCCCCCGGCTGTACCGACACATTTCCAACT 4015
Db 2340 GCGGTCCCCGCTTTTCAGAGACCACTCGTCCCTCCGGCGGTACCGACACATTTCCAGGT 2399
Qy 4016 GGCACATCTGCAGCTCTCTACTTGGCAGCGGCAAGACACCAAGTCCGGCTGTGCTATGC 4075

2400	Db	GGCCCATCTACACGCCCTTACTGCTAGCGCAAGAGCACTAAGGTGCGCGCTTCGCTATGC	2459
4076	Qy	AGCCCAAGGGTACAAGTGCTCGTCTCTGAAACCGTTCGTTCGCCACACCTTAAGGTGTTGG	4135
2460	Db	AGCCCAAGGGTAAAGTGCTGTTCTCTGAACCGTTCGTTCGCCACACCTTAAGGTTCGG	2519
4136	Qy	GGCGTATATGTCCAAGGCACACGGTATCGACCCCTAAATCATCAGAACTGGGGTAAAGACCAT	4195
2520	Db	GGCGTATATGTCFAAGGCACATGGTATCGACCCCTAAATCATCAGAAACCGGGGTAAAGACCAT	2579
4196	Qy	TACACGGGCGGCTCCATTAGCTACTCCACTATGCGAAGTTCTTTGCGCAGCGGTGGCTG	4255
2580	Db	CACACGGGTGCCCCCATCAGTACTCCACTATGCGAAGTTCTTTGCGCAGCGGTGGTTG	2639
4256	Qy	TTCTGGGGCGCCTATGACATCATATATGTGATGAGTGCACCTCAACTGACTCGACTAC	4315
2640	Db	CTCTGGGGCGCCTATGACATCATATATGTGATGAGTGCACCTCAACTGACTCGACCCAC	2699
4316	Qy	CATCTTGGGCATCGGCACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGCGCTCGTCTGT	4375
2700	Db	TATCTCGGCATCGGCACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGCGACTCGTCTGT	2759
4376	Qy	GCTCGCACCGCTACACCTCCGGGATCGGTTACGGTGGCCACACCCCAATATCGAGGAAT	4435
2760	Db	GCTCGCACCGCTACCGCTCCGGGATCGGTCACCGTGGCCACATCCAAAATCGAGGAGT	2819
4436	Qy	AGGCCTCTCCAAATCGAGAGATCCCTCTTATGCGAAAGCCATCCCAATTGAGGCCAT	4495
2820	Db	GGCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCAATCGAGACCAT	2879
4496	Qy	CAAGGGGGGAGGCATCTCAATTTCTGCCAATCCAAAGAAATGTGACAGCTCGCCGC	4555
2880	Db	CAAGGGGGGAGGCACCTCAATTTCTGCCAATCCAAAGAAATGTGATGAGCTCGCCGC	2939
4556	Qy	AAAGCTGACAGGCTTCGGACTGAACGCTGTAGCATATTACCGGGGCTTGTATGTGTCCT	4615
2940	Db	GAAGCTGTCCGGCTTCGGACTCAATGCTGTAGCATATTACCGGGGCTTGTATGTATCCCT	2999
4616	Qy	CATACCGCTATCGGAGACGTCGTTGCTGGCAACAGACGCTTAATGACGGTTTTCAC	4675
3000	Db	CATACCAACTAGCGGAGACGTCATTGTCGTAGAACGGACGCTTAATGACGGGCTTTAC	3059
4676	Qy	CGGCGATTTTGACTTCAGTGTACGACTGCAATACATGTGTCAACCCAGACAGTGCAGTTCAG	4735
3060	Db	CGGCGATTTTGACTTCAGTGTACGACTGCAATACATGTGTCAACCCAGACAGTGCAGTTCAG	3119
4736	Qy	CTTGGATCCCACTTCACTATTGAGACGACGCGTCCCCAAGACGCGGTGTGCGGCTC	4795
3120	Db	CCTGGACCCGACCTTCACTACATTGAGACGACCGCTGCCAAGACGCGGTGTGACGCTC	3179
4796	Qy	GCAACGGGAGGTAGACTGCGAGGGTAGGAGTGGCATCTACAGTFTTGTGACTCCAGG	4855
3180	Db	GCAGCGGAGGACGAGCTGGTATGGGACGAGTGGGCAATTTACAGTFTTGTGACTCCAGG	3239
4856	Qy	AGAACGGCTTCGGGCATGTTTCGATCTTTTCGGTCCCTGTGAGTGTCTATGACGCGGCTG	4915
3240	Db	AGAACGGCTTCGGGCATGTTTCGATCTTTTCGGTCCCTGTGAGTGTCTATGACGCGGCTG	3299
4916	Qy	TGCTTGGTATGAGCTCACGCCCGCTGAGACCTCGGTTAGGTTGCGGCTTACCTAATAC	4975
3300	Db	TGCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTGCGGCTTACCTAATAC	3359
4976	Qy	ACAGGGTTGCCCTCTGCGAGGACCATCTGGAGTCTGGGAGAGCTCTTCAAGSCCT	5035
3360	Db	ACCAGGGTTGCCCTCTGCGAGGACCATCTGGAGTCTGGGAGAGCTCTTCAAGSCCT	3419
5036	Qy	CACCCACATAGATGCCACTTCTGTGCCAGACTTAACAGGCAGGAGCAACTTTCCTTA	5095
3420	Db	CACCCACATAGAGCCCATTTCTGTGTCCAGACTTAAGGCAGGAGCAACTTTCCTTA	3479
5096	Qy	CCTGTTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACTCCATCGTGGGA	5155

QY 6236 TGAGACTGCTCTACGCCATGCTCCGGCTCGTGGCTAAAGGATGTTTGGGATTGGATATG 6295
DB |||||
QY 4620 CGAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTGGATATG 4679
DB |||||
QY 6296 CACGGTGTGACTGACTTCAAGACCTGGCTCAGTCCAAACTCTCGCCGCGGTACC GGG 6355
DB |||||
QY 4680 CACGGTGTGACTGATTTCAAGACCTGGCTCAGTCCAAAGCTCCTCGCCGCGATTGCCGGG 4739
QY 6356 AGTCCCTTTCCTGTCATGCCAAGCGGGGTACAAAGGAGTCTGGCGGGGGAGCGGCATCAT 6415
DB |||||
QY 4740 AGTCCCTTTCCTCATGTCAACGTGGGTACAAAGGAGTCTGGCGGGGGCGACGGGCATCAT 4799
QY 6416 GCAAAACCACTGCCATGCGGAGCAGATGCCGGAATGTCAAAAACGGTTCATCAG 6475
DB |||||
QY 4800 GCAAAACCACTGCCATGCGGAGCAGATCACCGGACATGTGA AAAACGGTTCATCAG 4859
QY 6476 GATCGTAGGGCTAGAACCTGCAGCAACACGTGGCACGGAACGTTTCCCATCAACGCGATA 6535
DB |||||
QY 4860 GATCGTGGGGCTAGGACCTGTAGTAACACGTGGCATGGAAACATTTCCCATTAACGCGTA 4919
QY 6536 CACCAAGGACCTTGCACACCTTCCCGCGGCCCAACTATTCAGGGGCGCTATGCGGGT 6595
DB |||||
QY 4920 CACCAAGGCGCTGCACGCCCTCCCGCGGCCAAATTAATTCAGGGCGCTGTGGCGGT 4979
QY 6596 GCTGCTGAGGAGTACGTGGAGGTACGCGTGTGGGGATTTCCACTACGTACCGGGCAT 6655
DB |||||
QY 4980 GCTGCTGAGGAGTACGTGGAGGTACGCGGTGGGGATTTCCACTACGTACCGGGCAT 5039
QY 6656 GACCACTGACAACTAAAGTGCCTATGCCAGGTTCGGGCCCGCGAATTTCTTCAACGAGGT 6715
DB |||||
QY 5040 GACCACTGACAACTAAAGTGCCTGTTCAGGTTCCGGCCCCCGAATTTCTTCAACAGAGT 5099
QY 6716 GGATGAGTGGGTTGCCACAGTACGCTCCGGGTGCAAACTCTTCAAGGGAGACGT 6775
DB |||||
QY 5100 GGATGGGTGCGGTTGACACAGGTACGCTCCAGCGTGC AAAACCCCTCTTACGGGAGGAGT 5159
QY 6776 CACGTTCCAGGTGCGGCTCAACCAATCTTGGTCCGGTCCGAGCTCCCATCGAGCCCGA 6835
DB |||||
QY 5160 CAACTTCTTGGTGGGCTCAATCAATACCTGTGGTTCAGAGCTCCCATGCGAGCCCGA 5219
QY 6836 ACCGACGTAAACAGTCTTACTTCCATGCTCACCGATCCCTCCCACTTACACGACAGAC 6895
DB |||||
QY 5220 ACCGACGTAGCAGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACCGCGGAGAC 5279
QY 6896 GGTAAAGCTAGGCTGGCTAGAGGGTCTCCCGCTCTTTAGCAGCTCATCAGCTAGCCA 6955
DB |||||
QY 5280 GGTAAAGCTAGGCTGGCTGGCAGGGGATCTCCCGCTCTTTGGCGAGCTCATCAGCTATCCA 5339
QY 6956 GTTGTCTGCGCTTCTTTGAAGCGACATGCATACCCACCATGACTCCCGCGGAGCTGA 7015
DB |||||
QY 5340 GCTGTCTGCGCTTCTTTGAAGGCAATGCACTACCCGTATGACTCCCGGAGCGCTGA 5399
QY 7016 CTTCACTGAGGCCAACCTCTTGTGGCGGAGAGATGGCGGAAAACATCACTCGCGTGA 7075
DB |||||
QY 5400 CTTCACTGAGGCCAACCTCTGTGGCGGAGAGATGGCGGGAACATCAACCGCGTGA 5459
QY 7076 GTCAGAGAAATAAGGTAGTAATTTCTGAATCTTTTGAACCGCTTTCACCGGAGGGGATGA 7135
DB |||||
QY 5460 GTCAGAAAAATAAGGTAGTAATTTTGGACTCTTTTCGAGCGGCTCCCAAGCGGAGGATGA 5519
QY 7136 GAGGAGATATCCGTCGCGCGGAGATCTGGGAAAATCCAGGAAGTTCCTCCCTCAGCGTT 7195
DB |||||
QY 5520 GAGGGAAGTATCCGTTTCCGCGGAGATCTGGGAGGTTCAGGAAATTTCCCTCGAGCGAT 5579
QY 7196 GCCATATGGGACGCGCGACTACAACTCTCCACTGTAGAGTCTCGGAAAGACCCGGA 7255
DB |||||
QY 5580 GCCATATGGGACGCGCGGATTACAACTCCACTGTTAGAGTCTCTGGAAGGACCCGGA 5639
QY 7256 CTAAGTCCCTCGGTTGGTACAGGATGCCCATTTGCCACTTACAAAGGCTCTTCCCAATACC 7315
DB |||||
QY 5640 CTAAGTCCCTCCAGTGGTACACGGGTGATTCATTTGCCGCTGCGCAAGGCGCTCCGATACC 5699

QY 7316 ACCTCCACGGAAGAAGGACGGTGTCTTGACAGAAATCCAAATGTGTCTTCTGCTTGGC 7375
DB |||||
QY 5700 ACCTCCACGGAAGAAGGACGGTGTCTTGACAGAAATCCAAATGTGTCTTCTGCTTGGC 5759
DB |||||
QY 7376 GGAGCTCGCCACTAAGACCTTCCGTAGCTCCGGATCGTCCGCCGTTGATAGCGGACGGC 7435
DB |||||
QY 5760 GGAGCTCGCCACAAAGACCTTCCGCAGCTCCGAAATCGTCCGCCGTCGACAGCGCACGGC 5819
QY 7436 GACCGCTTCTCTCACCTCGGCTCCGACGAGCGGTGACAAAGATCCCGACGTTGAGTCGTA 7495
DB |||||
QY 5820 AACGCTCTCTCTGACACAGCCCTCCGACGAGCGGACCGGGATCCGACGTTGAGTCGTA 5879
QY 7496 CTCCTCCATGCCCCCTTGAAGGGGAGCCGGGGACCCCGATCTCAGCGACGGGTCTTG 7555
DB |||||
QY 5880 CTCCTCCATGCCCCCTTGAGGGGAGCGGGGGATCCCGATCTCAGCGACGGGTCTTG 5939
QY 7556 GTCTACCGTGTAGGAGGCTAGTGAGGATGTGCTGTGCTGCTCAATGTCTTATACGTG 7615
DB |||||
QY 5940 GTCTACCGTAAAGCAGGAGGCTAGTGAGGACGTGCTGTGCTGCTCGATGTCTACACATG 5999
QY 7616 GACAGGCGCTGATCAGCCATCGCTGCGGAGGAAAGTAAGCTGCCCATCAACCCGTT 7675
DB |||||
QY 6000 GACAGGCGCTGATCAGCCATCGCTGCGGAGGAAACCAAGCTGCCCATCAATGCACT 6059
QY 7676 GAGCAACTTTTGTGCTCACCAACATGCTCTAGCCCAACAACATCCCGCAGCGCAAG 7735
DB |||||
QY 6060 GAGCAACTTTTGTCCGTCCGTACACAACTTGTCTATGTACAACTCTCGCAGCGCAAG 6119
QY 7736 CTTCCGGGAGAAGGTACCTTTGACAGATGCGAAGTCTCTGGATGATCATATTACCGGA 7795
DB |||||
QY 6120 CTTCCGGGAGAAGGTACCTTTGACAGATGCGAAGGTCTCAGAGTCTCGACGACCACTACCGGA 6179
QY 7796 CGTACTCAAGGAGATGAAGCGAGGCGTCCACAGTTAAAGCTTAAGCTTCTATCATAGA 7855
DB |||||
QY 6180 CGTGTCAAGGAGATGAAGCGAAGGCGTCCACAGTTAAAGCTTAAGCTTCTATCGTGA 6239
QY 7856 GGAGCCTTGCAAGCTGACGCCCCACATTTCCGCGCAAAATCCAAATTTGGCTATGGGCGAAA 7915
DB |||||
QY 6240 GGAAGCCTGTAACTGACGCCCCACATTTCCGCCAGATCTAAATTTGGCTATGGGCGAAA 6299
QY 7916 GGAGCTCCGGAACTTATCCAGCAGGCGGTTAAACACATCCGCTCCGTTGGGAGGACTT 7975
DB |||||
QY 6300 GGAAGCTCCGGAACTTATCCAGCAGGCGGTTAAACACATCCGCTCCGTTGGGAGGACTT 6359
QY 7976 GCTGGAAGACACTGAACACCAATTTGACACCACTCATGCGCAAAAGTGAAGTCTTCTG 8035
DB |||||
QY 6360 GCTGGAAGACACTGAGACACCAATTTGACACCACTCATGCGCAAAAGTGAAGTCTTCTG 6419
QY 8036 CGTCCAAACGAGAGAGGAGCGCCGACGCTCGCTTATCGTATTTCCACAGACCTGGG 8095
DB |||||
QY 6420 CGTCCAAACGAGAGAGGAGCGCGCAGCAGCTCGCTTATCGTATTTCCAGATTTGGG 6479
QY 8096 AGTTGCTGTATGCGAGAGATGGCCCTTTTACGAGTGTCTCCACCTTCTCAGGCGGT 8155
DB |||||
QY 6480 GGTTCGTGTGCGAGAAATGGCCCTTTAGCATGTGCTCTCCACCTCCCTCAGGCGGT 6539
QY 8156 GATGGCTCTCATACGATTTCAATCTCCCGCAAGCAGCGGTCGAGTTCCTGCTGAA 8215
DB |||||
QY 6540 GATGGCTCTCATACGATTTCCAACTCTCTCTGGACAGCGGTGCGAGTTCCTGCTGAA 6599
QY 8216 TACCTGGAATCAAAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGACTC 8275
DB |||||
QY 6600 TCCCTGGAAGGAGAAATGCCCTTATGGGCTTCGCTATGACACCCGCTGTTTGACTC 6659
QY 8276 AACGCTCATGAGAGTGAATTTGAGAGTCAATTTTCAATTTCAATTTGAGCTTGGC 8335
DB |||||
QY 6660 AACGCTCATGAGAGTGAATTTGAGAGTCAATTTTCAATTTTCAATTTGAGCTTGGC 6719
QY 8336 CCCCAGGCGCAGACGAGCCATAAGTTCGCTCACAGCGGCTTTACATCGGGGTCCTC 8395
DB |||||
QY 6720 CCCCAGGCGCAGACGAGCCATAAGTTCGCTCACAGCGGCTTTTACATCGGGGTCCTC 6779
QY 8396 GACTAACTCAAAAGGCGAGAACTGCGGTTTATCGCGGTCGCGCGCAAGTGGCGTGTGAC 8455

[illegible]

Db	3360	ACACGGTGTCCCGTCTGCCAGACCACTGGAGTTCCTGGGAGAGCGTCTTTACAGCGCT	3419
Qy	5036	CACCCACATAGATGCCCACTTCTCTGCCAGACTAAACAGGCGAGGAGACAACATTTTCCCTTA	5095
Db	3420	CACCCACATAGACGCCATTTCTTGTCCAGACTAAGCAGGCGAGGAGACAATTCCTCCCTA	3479
Qy	5096	CTGTGGGCATATCAAGCTACAGTGTGGCCAGGGCTCAAGCTCCACTTCATCTGTGGGA	5155
Db	3480	CCTGGTAGCATACAGCTACGGTACGGTGGCCAGGGCTCAGGCTCCACCTCCATCGTGGGA	3539
Qy	5156	CCAAATGTGAAGTGTCTCATACGGCTGAACCTACACTGCACGGGCCAACCCCTGCT	5215
Db	3540	CCAAATGTGAAGTGTCTCATACGGCTAAAGCTCAGCTGCACGGGCCAACGCCCTGCT	3599
Qy	5216	GTATAGCTAGGAGCGCTCCAAATAGGTCATCTTCACACACCCCACTAAATACAT	5275
Db	3600	GTATAGCTGGAGCGCTTCAAACGAGGTTACTACACACACCCCATACCAATACAT	3659
Qy	5276	CATGGCATGATGTCCGCTGACCTGGAGTCTGTCACTAGCACCTGGGTGTGTAGCGG	5335
Db	3660	CATGGCATGATGTCCGCTGACCTGGAGTCTGTCACTAGCACCTGGGTGTGTAGCGG	3719
Qy	5336	AGTCTTGCAGCTTTGGCCGCATCTACCTGTGACGACAGGCAGTGTGTCTATTGTGGCAG	5395
Db	3720	AGTCTTAGCAGCTCTGGCCGCTATTTCCTTGAACAACAGGCAGCGTGTCTATTGTGGCAG	3779
Qy	5396	GATCATCTTCTCGGGAAGCAGCTGTCTGTTCGCCGACAGGGAAGTCTCTACACAGAGTT	5455
Db	3780	GATCATCTTCTCGGGAAGCAGCTGTTCGCCGACATCTCCGACAGGGAAGTCTTTACCGGAGTT	3839
Qy	5456	CGATGAGATGGAAGAGTGCCTCACAACTTCCTTATCATCGACAGGGAATCAGCTCTGC	5515
Db	3840	CGATGAGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATCAGCTCTGC	3899
Qy	5516	CGACCAATTCAGCAAAAGCGCTCGGTTGTTCGCAACGGCCACACGAGGAGGC	5575
Db	3900	CGAACAAATTCAGCAAAAGCAATCGGTTGTCTGCAAAACGCCACCAAGCAAGCGGAGGC	3959
Qy	5576	TGCTGCTCCCGTGTGAGTCCAAAGTGGCAGCCCTTGAGACCTTCTGGGCGCAAGACAT	5635
Db	3960	TGCTGCTCCCGTGTGAGTCCAAAGTGGCAGCCCTTGAAGCCTTCTGGGCGCAAGCATAT	4019
Qy	5636	GTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTCGAACC	5695
Db	4020	GTGGAATTTTCATCAGCGGGATACAAATATTAGCAGGCTGTCTCCACTCTGCTTGGCAACC	4079
Qy	5696	CGCGATAGCATCATGTAGGCAATTAAGCTTCTATCACTAGCCGCTCACCACCCAAA	5755
Db	4080	CGCGATAGCATCATGTATGGCATTTACAGCTTCTATCACTAGCCGCTCACCACCCAAA	4139
Qy	5756	CACCTCCTGTTTAACTTTTGGGGGATCGGTGCTGCCCACTCGCTCTCTCCACGCG	5815
Db	4140	TACCTCCTGTTTAACTTTTGGGGGATCGGTGCGCCCACTTCTCTCCACGCG	4199
Qy	5816	TGCGTAGCTTTCTGTGGCGCCCGGCATCTGCGGAGCGGCTTGTGGCAGCATAGGCTTGG	5875
Db	4200	TGCTTCTGCTTTCTGTAGCGCCGCGCATCGCTGGAGCGGCTTGTGGCAGCATAGGCTTGG	4259
Qy	5876	GAAGGTGCTGTGACATCTTTGGCGGCTATGGGGCAGGGGTAGCGCGCGCATCTGTGGC	5935
Db	4260	GAAGGTGCTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCTGTGGC	4319
Qy	5936	CTTTAAGTTCATAGCGCGGAGTGCCCTCCACCGAGGACTGTGTCAACTTACTCCCTGC	5995
Db	4320	CTTTAAGTTCATAGCGCGGAGATGCCCTCCACCGAGGACTGTGTCAACTTACTCCCTGC	4379
Qy	5996	CATCCTCTCTCTGGTGCCTGTGTCTGGGGTCTGTGGCGAGCAATACTTCGTTCGGCA	6055
Db	4380	TATCCTCTCCCTGCGCCCTAGTCTGTGGGGTCTGTGGCGAGCATCTTCGTTCGGCA	4439
Qy	6056	CGTGGGCCCGGAGAGGGGGCTGTGCAATGGAACCGGCTGTATAGCGTTTCGTTCGCG	6115
Db	4440	CGTGGGCCCGGAGAGGGGGCTGTGCAATGGAACCGGCTGTATAGCGTTTCGTTCGCG	4499

Qy	6116	GGTAAACCAAGTCTCCCTACGCACTATGTCCTGAGAGCGACGCTGACAGCTGTAC	6175	1796	GCCCATATGGGACGCGCGGACTACAATCTCCACTGTAGAGTCTCTGAAAGGACCCGGA	7255
Db	4500	GGGTAAACCAAGTCTCCCTACGCACTATGTCCTGAGAGCGACGCTGACAGCTGTAC	4559	5580	GCCCATATGGGACGCGCGGANTTAAACCTTCACTGTTAGAGTCTCTGAAAGGACCCGGA	5639
Qy	6176	TCAGATCCTCTAGCTTACCATCACTCACTGCTGAAGCGCTCCACAGTGGATTAA	6235	7256	CTACGTCCTCTCGGTGGTACACGGATGCCCATTGGCCACTTACCAAGCTCTCTCAATACC	7315
Db	4560	TCAGATCCTCTAGCTTACCATCACTCACTGCTGAAGCGCTCCACAGTGGATTAA	4619	5640	CTACGTCCTCTCGGTGGTACACGGGTGTCAATGGCGCTGCGCAAGGCGCTCCGATACC	5699
Qy	6236	TGAGGACTGCTACAGCCATGCTCCGGCTCGTGGCTAAGGATGTTTGGATTGGATAG	6295	7316	ACCTCACCGAGAAAGAGACGGTTGCTCTGACAGAAATCCAAATGCTCTCTGCGCTTGGC	7375
Db	4620	CGAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATAG	4679	5700	ACCTCACCGAGAAAGAGACGGTTGCTCTGACAGAAATCTACCGTGTCTCTGCGCTTGGC	5759
Qy	6296	CACGGTGTGACTGACTTTCAAGACCTGGCTCCAGTCCAAACTCTGCGCGGGTTACCGGG	6355	7376	GGAGCTCGGCACCTAAGACCTTCGGTAGCTCCGGATCGTGGCGGTGTATAGCGGCGACGGC	7435
Db	4680	CACGGTGTGACTGACTTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGGGTTACCGGG	4739	5760	GGAGCTCGGCACAAAGACCTTCGGTAGCTCCGGATCGTGGCGGTGTATAGCGGCGACGGC	5819
Qy	6356	AGTCCCTTCTGCTGATGCAACGCGGTACAGGGAGTCTGGCGGGGACGGCATCAT	6415	7436	GACCGCCCTTCTGACCTCGGCCCTCCGACGAGCGGTGACAAAGGATCCGACGTTGAGTCGTA	7495
Db	4740	AGTCCCTTCTGCTGATGCAACGCGGTACAGGGAGTCTGGCGGGGACGGCATCAT	4799	5820	NACGGCTCTCTCTGACCGCTCCGACGAGCGGATCCGAGATCCGACGTTGAGTCGTA	5879
Qy	6416	GCAACCACTGCCATGCGGAGCACAGATCGCGGACATGTCAAAAACGTTCCATGAG	6475	7496	CTCCTCATGCCCCCTTGAAGGGGAGCGGGGACCCCGATCTCAGCGACGGGCTTTCG	7555
Db	4800	GCAACCACTGCCATGCGGAGCACAGATCGCGGACATGTGAAAACGTTCCATGAG	4859	5880	CTCCTCATGCCCCCTTGAAGGGGAGCGGGGATCCCGATCTCAGCGACGGGCTTTCG	5939
Qy	6476	GATCGTAGGCTAGAACCTTGCAGCAACACGTGGCAGCGAAGCTTCCCATCAACGCAITA	6535	7556	GTCTACCGGTGAGTAGGAGCTAGTCAGGATGTCTGCTGCTCAATGTCTTATACGTG	7615
Db	4860	GATCGTAGGCTAGAACCTTGCAGCAACACGTGGCAGCGAAGCTTCCCATCAACGCGTA	4919	5940	GTCTACCGTAAGCGAGAGCTAGTCAGGAGCTGCTGCTGCTCAATGTCTTATACATG	5999
Qy	6536	CACACGGGACCTTGCACACCTCCCGGCGCCCACTATTTCAGGGCGCTATGGCGGT	6595	7616	GACAGCGCCCTGATCAGCCATGCGCTGCGGAGGAAGTAAGCTGCCCATCAACCCGTT	7675
Db	4920	CACACGGGACCTTGCACACCTCCCGGCGCCCACTATTTCAGGGCGCTATGGCGGT	4979	6000	GACAGGGCCCTGATCAGCCATGCGCTGCGGAGGAACCAAGCTGCCCATCAATGCACT	6059
Qy	6596	GGCTGCTGAGGAGTACGTGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGCAT	6655	7676	GAGCAACTCTTTGCTGGTCAACAAACATGCTTACGCCACAACTCCCGACGCGCAAG	7735
Db	5039	GGCTGCTGAGGAGTACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGCAT	5099	6060	GAGCAACTCTTTGCTGGTCAACAACTTGGTCTATGCTTACAACTCTCGACGCGCAAG	6119
Qy	6656	GACCACTGACAACTAAAGTSCCATGCCAGTTCCGGCCCCCGAAATCTTTCACGGAGGT	6715	7736	CTCCCGGACAGAAAGTCACTTTTGACAGATTGCAAGTCTCTGATGATCATTAACCGGGA	7795
Db	5040	GACCACTGACAACTAAAGTSCCATGCCAGTTCCGGCCCCCGAAATCTTTCACAGAGT	5099	6120	CTCGGCGAGAAAGTCACTTTTGACAGCTGCAAGTCTTGGACGACCACTACCGGGA	6179
Qy	6716	GGATGAGTGGGTTGACAGGATAGCTCCGGCTGCAAACTCTTTCACGGGAGGAGT	6775	7796	CGTACTCAAGAGATGAAGCGGAGGCTCCACAGTTAAGGCTAAGCTTATCCGTTGA	7855
Db	5100	GGATGAGTGGGTTGACAGGATAGCTCCGGCTGCAAACTCTTTCACGGGAGGAGT	5159	6180	CGTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTTAAGCTTAACTTCTATCCGTTGA	6239
Qy	6776	CACGTTCCAGTCCGGCTCAACCAATCTTGGTGGGTCGAGCTCCCATCCGAGCGCGA	6835	7856	GGAGGCTGCAAGCTGACGCCCCCACTTGGCGCAATCCAAATTTGGCTATGGGGCAAA	7915
Db	5160	CACATCTCTGGTGGGCTCAATCAATCTTGGTGGGTCACAGCTCCCATCCGAGCGCGA	5219	6240	GGAGGCTGTAAGCTGACGCCCCCACTTGGCGCAGATCTAAATTTGGCTATGGGGCAAA	6299
Qy	6836	ACCGGAGTAAAGTGTCTTCCATGCTCACCGATCCCTCCACATTAACAGCAGAGAC	6895	7916	GGAGCTCCGGAACCTTATCAGCAGCGGCGGTTAACCATCCGCTCGGTGGGAGGACTT	7975
Db	5220	ACCGGAGTAAAGTGTCTTCCATGCTCACCGATCCCTCCACATTAACAGCAGAGAC	5279	6300	GGAGCTCCGGAACCTTATCAGCAGCGGCGGTTAACCATCCGCTCGGTGGGAGGACTT	6359
Qy	6896	GGTAAAGCTAGGCTGAGGAGTCTCCCTCTTTTACGAGCTCATCAGCTAGGCA	6955	7976	GCTGGAAGACACTGAAACACCAATTGACACCACTCATGGCAAAAAGTGAAGGTTTCTG	8035
Db	5280	GGTAAAGCTAGGCTGAGGAGTCTCCCTCTTTTACGAGCTCATCAGCTAGGCA	5339	6360	GCTGGAAGACACTGAGACACCAATTGACACCACTCATGGCAAAAAGTGAAGGTTTCTG	6419
Qy	6956	GTGTCTCGGCTTTTGAAGGACATGACATACCACTGCTCCCGGAGCGTGA	7015	8036	CGTCCAAACAGAGAGGAGCGCGCAAGCGAGCTCGCTTATTCGTATTTCCAGACCTTGGG	8095
Db	5340	GTGTCTCGGCTTTTGAAGGACATGACATACCACTGCTCCCGGAGCGTGA	5399	6420	CGTCCAAACAGAGAGGAGCGCGCAAGCGAGCTCGCTTATTCGTATTTCCAGATTTGGG	6479
Qy	7016	CCTCATCAGGCGCAACCTTGTGGCGCAGGAGATGGCGGAAACATCACTCCGCTGGA	7075	8096	AGTTCTGTATGCGAAGAGATGGCCCTTACGAGCTGCTCCACCTTCTCTCAGGCGGT	8155
Db	5400	CCTCATCAGGCGCAACCTTGTGGCGCAGGAGATGGCGGAAACATCAACCGCTGGA	5459	6480	GGTTCGTGTGCGGAGAAAATGGCCCTTACGATGTGGTCTCTCCACCTCTCAGGCGGT	6539
Qy	7076	GTCAGAGAAATAGGTAGTAAATTTGAGCTTTTTCGAACCGCTTTCACGGGAGGAGTGA	7135	8156	GATGGGCTCTCATACAGATTTCAATCTCCCAAGCAGCGGCTCGAGTCTCTGTTGAA	8215
Db	5460	GTCAGAAATAGGTAGTAAATTTGAGCTTTTTCGAACCGCTTTCACGGGAGGAGTGA	5519	6540	GATGGGCTCTCTATACCGATTTCAATCTCTCTGGAACGCGGCTCGAGTCTCTGTTGAA	6599
Qy	7136	GAGGAGATATCCGTGCGGCGGAGATCGTGGGAAATTCAGGAAAGTTCCCTCAGCGTT	7195	8216	TACCTGGAATCAAGAAATGCGCTTATGGGCTTCTCATATGACACCGCTGTTTGAATC	8275
Db	5520	GAGGAGATATCCGTGCGGCGGAGATCGTGGGAAATTCAGGAAAGTTCCCTCAGCGAT	5579	6600	TGCTGGAAGCGGAGAAATGCGCTTATGGGCTTTCGCATATGACACCGCTGTTTGAATC	6659
Qy				8276	AACGGTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTTACCAATGTTGTGACTTGGC	8335

Db 2160 GGGCGACAGCGGGAGGCTACTCTCCCGCCAGGCCGCTCTCTACTTTGAAGGGCTCTTC 2219
QY 3836 GGGTGGTCCATTGCTTTGGCCCTTCGGGGACGFCGCTGGGGCTCTTCGGGGCTGCTGTGTG 3895
Db 2220 GGGCGGTCACTGCTCTGCCCCCTCGGGGCACGCTGTGGGCACTCTTCGGGGCTGCCGTGTG 2279
QY 3896 CACCCGGGGGTTCGAAAGCGGTGACCTTTCATACCCGTTGAGTCTATGGAACCTACCAT 3955
Db 2280 CACCCGAGGGTTGCGAAGCGGTGACCTTTGTACCCGTCGAGTCTATGGAACCACTAT 2339
QY 3956 GCGGTCTCCGGTCTTACAGACAACTCNAACCCCGCGGTGTACCCGACGACATTTCCAAGT 4015
Db 2340 GCGGTCCCAGGTCTTACAGGAACTCTGTCCTGAAACCCGTCGTCGCCGCCACCCCTAGGTTTCGG 2399
QY 4016 GGCACATCTCAGCGCTCTACTGGCAGCGGCAAGAGCACAAAGTGC CGGCTCGGTATGC 4075
Db 2400 GGGCCATCTACAGCCCTACTCTGTTAGGCGGCAAGACACTAAGTGC CGGCTCGGTATGC 2459
QY 4076 AGCCCAAGGGTACAAGGTGCTCGTCTGAAACCCGTCGTTGCGGCCACCTTATAGGTTTGG 4135
Db 2460 AGCCCAAGGGTATAAGGTGCTTGTCTGAAACCCGTCGTCGCCGCCACCCCTAGGTTTCGG 2519
QY 4136 GGGTATATGTTCAAGGCCACCGTATCGACCCCTTAACATCAGAACTGGGGTAAAGACCAT 4195
Db 2520 GGGGTATATGTTCAAGGCCACCGTATCGACCCCTTAACATCAGAAACCGGGTAAAGACCAT 2579
QY 4196 TACCAGGGGGTCCAACTAGCTACTCCACTATGGCAAGTCTCTTCCGACCGGTGCTG 4255
Db 2580 CACACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTCTCTTCCGACCGGTGCTG 2639
QY 4256 TTCTGGGGGGCCCTATGACATCATATATGATGATGAGTGCCTCAACTCACTGACCTGACATC 4315
Db 2640 CTCTGGGGGGCCCTATGACATCATATATGATGATGAGTGCCTCAACTCACTGACCTGACATC 2699
QY 4316 CATCTTGGGCATCGGCACAGTCTGGACCAAGCGGACACGGCTGGAGCGGGCTCGTCTG 4375
Db 2700 TATCTGGGCATCGGCACAGTCTCGGACCAAGCGGAGACGGCTGGAGCGGCTCGTCTG 2759
QY 4376 GCTCGCCACCGCTACACTCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAAAT 4435
Db 2760 GCTCGCCACCGCTACCGCTCGGGATCGGTTACCGTGCCACACCCCAATATCGAGGAAAT 2819
QY 4436 AGGCTCTCCAAATGGAGAGATCCCGCTCTATGGCAAGGCCATCCCAATGAGGCCAT 4495
Db 2820 GGCTCTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGGCCATCCCAATGAGGCCAT 2879
QY 4496 CAAGGGGGGAGGCATCTCATTTTCTGCTCATPTCCAAAGAAATGTGACGAGCTCGCGGC 4555
Db 2880 CAAGGGGGGAGGCACCTCATTTTCTGCTCATPTCCAAAGAAATGTGATGAGCTCGCGGC 2939
QY 4556 AAAGCTGACAGGCTTCGACGTGACGCTGTAGCATATACCGGGCCCTTGTATGTCGCT 4615
Db 2940 GAAAGCTGTCCGGCCCTCGGACTCAATGTGTGTAGCATATACCGGGCCCTTGTATGTCGCT 2999
QY 4616 CATACCGCTATCGAGACGCTGTGTGTCGGCAACAGACGCTCTAATGACGGGTTTCAC 4675
Db 3000 CATACCAACTAGCGGAGACGCTATGTGTGAGCAACGACGCTCTAATGACGGGCTTAC 3059
QY 4676 CGGCGAATTTGACTCAGTGATCGACTGCAATACATGTGTCAACGACAGCTGCACTTCAG 4735
Db 3060 CGGCGAATTTGACTCAGTGATCGACTGCAATACATGTGTCAACGACAGCTGCACTTCAG 3119
QY 4736 CTTGGATCCCACTTACCATTTAGACAGACCGGTGCCCCAAGACGGGTGTCGCGCTC 4795
Db 3120 CTTGACCCGACCTTACCATTTAGACAGACCGGTGCCCCAAGACGGGTGTCACGCTC 3179
QY 4796 GCACCGGAGGTAGAACTGCGAGGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGG 4855
Db 3180 GCAGCGGAGGAGGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGG 3239
QY 4856 AGACGGCCCTCGGGCATGTTCTGATTTCTCGGTCTCTGTGAGTGTCTATGACGGGGCTG 4915

Db 3240 AGAACGGCCCTCGGGCATGTTCTGATTCTCTCGGTTCGTGCGAGTGTCTATGACGCGGGCTG 3299
QY 4916 TGCCTTGGTATAGCTCAACGCCCGCTGAGACCTCGGTAGTTCGGGGCTTACTTAATAC 4975
Db 3300 TGCCTTGGTACGAGCTCAACGCCCGCGGAGACCTCAGTTAGTTCGGGGCTTACTTAATAC 3359
QY 4976 ACCAGGGTTGCCGCTGCGCAGGACCATCTGAGATTCTGGGAGAGCGTCTTTCACAGGCCT 5035
Db 3360 ACAGGGTTGCCGCTGCGCAGGACCATCTGAGATTCTGGGAGAGCGTCTTTCACAGGCCT 3419
QY 5036 CACCCACATAGATGCCACTTCTGTCAGACTAAACAGGCGAGGAGACAACTTTCCTTA 5095
Db 3420 CACCCACATAGAGCGCCCACTTCTGTCAGACTAAAGCAGGCGAGGAGACAACTTTCCTTA 3479
QY 5096 CTTGTTGGCATATCAAGCTACAGTGTGCGCGAGGCTCAAGCTCCACCTCCATCTGTTGGGA 5155
Db 3480 CTTGTTGAGCATACAGGCTACGTTGTCGCGCAGGCTCAAGGCTCCACCTCCATCTGTTGGGA 3539
QY 5156 CCAATGTGGAAGTGTCTCATACGGCTGAAACCTCATCTGACGCGGCCAACACCCCTGCT 5215
Db 3540 CCAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGACGCGGCCAACACCCCTGCT 3599
QY 5216 GTATAGGCTAGGACCGCTCCAAATGAGGTCACTCTCACACACCCCATAACTAAATACAT 5275
Db 3600 GTATAGGCTAGGACCGCTCCAAACCGAGGTTACTACACACACCCCATAACTAAATACAT 3659
QY 5276 CATGGCATGATGTCGGCTGACCTGAGGTGCTCACTAGCACCTGGGTGCTGTAGCGG 5335
Db 3660 CATGGCATGATGTCGGCTGACCTGAGGTGCTCACGAGCACCTGGGTGCTGTAGCGG 3719
QY 5336 AGTCCTTGCAGCTTTCGGCGCATCTGCTGACGACAGGCGAGTGTGCTCACTGTGGGCG 5395
Db 3720 AGTCCTTGCAGCTTTCGGCGCATCTGCTGACGACAGGCGAGTGTGCTCACTGTGGGCG 3779
QY 5396 GATCATCTTGTTCGGGAAGCCAGCTGTGTTCCCGACAGGGAAGTCTCTACACGAGGTT 5455
Db 3780 GATCATCTTGTTCGGGAAGCCGCGCATCATTTCCCGACAGGGAAGTCTCTACCGGAGTT 3839
QY 5456 CGATGAGATGGAAGAGTGTGCTCAACATTTCTTATCATCGAGCAGGGAATGACGTGCG 5515
Db 3840 CGATGAGATGGAAGAGTGTGCTCAACATTTCTTATCATCGAAGCAGGGAATGACGTGCG 3899
QY 5516 CGAGCAATTCAGCAAAAGCGCTCGGTTGTTGCAAAAGCGCCACCAAGCGGAGGC 5575
Db 3900 CGAACAATTCAGCAAAAGCGCAATTCGGGTTGCTGCGCAACAGCCACCAAGCGGAGGC 3959
QY 5576 TGCTGCTCCGCTGGTGAGTCCAAAGTGGCGAGCCCTTGAGACCTTTCTGGGCGAAGCAGAT 5635
Db 3960 TGCTGCTCCGCTGGTGGAATCCAAAGTGGCGGACCCCTCGAAGCCCTTCTGGGCGAAGCAGAT 4019
QY 5636 GTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGCGGAAACCC 5695
Db 4020 GTGGAATTTTCATCAGCGGATACAAATATTTAGCAGGCTTGTCCACTCTGCTGCGCAACCC 4079
QY 5696 CGGATAGCATCATTTGATGCAATTTACAGCTTCTATCACTAGCGGCTCAACCCCAAAA 5755
Db 4080 CGGATAGCATCATTTGATGCAATTTACAGCTTCTATCACTAGCGGCTCAACCCCAAAA 4139
QY 5756 CACCTCCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCCCAACTCGCTCTCTCCAGGCG 5815
Db 4140 TACCTCCTCTGTTTAAACATCTTGGGGGATGGGTGGCGGCCCAACTCTGCTCTCCAGGCG 4199
QY 5816 TGCCTAGCTTTCGTTGGCGCGCGGCATCGCGGAGCGGCTGTGCGCAGCATAGGCTTGG 5875
Db 4200 TGCCTTCTGCTTTCGTTAGCGCGCGGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGG 4259
QY 5876 GAAAGTGTCTGTGACATCTTGGCGGCTATGCGGCGAGGGGTAGCGCGGCACCTCGTGGC 5935
Db 4260 GAAAGTGTCTGTGACATCTTGGCGGCTATGCGGCGAGGGGTGCGAGCGCGCTCGTGGC 4319
QY 5936 CTTTAAAGGTCTAGCGCGGAGGTCCTCCACCGAGGACCTGTTCAACTTACTTCCCTGCG 5995
Db 4320 CTTTAAAGGTCTAGCGCGGAGATGTCCTCCACCGAGGACCTGTTTAACTTACTTCCCTGCG 4379


```
Db 6540 GATGGGCTCTTCAATCGGATTCCTCTCGACAGCGGGTCAAGTTCTCGTGAA 6599
Qy 8216 TACTTGAAATCAAGAAATGCGCTATGGGCTTCTCATATGACACCGCGTGTGTTGACTC 8275
Db 6600 TGCCTGGAAGCGAAGAAATGCGCTATGGGCTTCGCATATGACACCGCGTGTGTTGACTC 6659
Qy 8276 AACGGTCACTGAGAGTGAATTCGTTGAGGAGTCAATTTACCAATGTTGACTTGGC 8335
Db 6660 AACGGTCACTGAGAAATGACATCGTGTGAGGAGTCAATTTACCAATGTTGACTTGGC 6719
Qy 8336 CCCGAGGCCAGACAGCCATAGTTCGCTACAGAGCGCTTTACATCGGGGCTCCCT 8395
Db 6720 CCCGAGGCCAGACAGCCATAGTTCGCTACAGAGCGCTTTACATCGGGGCGCCCT 6779
Qy 8396 GACTAATCTAAAGGGCAGAACTGCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGAC 8455
Db 6780 GACTAATCTAAAGGGCAGAACTGCGGCTATCGCGGTGCGCGCAAGTGGTGTGAC 6839
Qy 8456 GACTAGTCGGGTAATACCTCATGTTACTTGAAGGCCACTGCGAGCCGTGCGAGTGC 8515
Db 6840 GACCAGTCGGTAATACCTCATGTTACTTGAAGGCCGCTGCGGCTGTGCGAGTGC 6899
Qy 8516 AAGCTCCAGACTGCGAGTCTCGTGAACGGAGACGCTTGTGCTTATCTGTGAAG 8575
Db 6900 AAGCTCCAGACTGCGAGTCTCGTGAACGGAGACGCTTGTGCTTATCTGTGAAG 6959
Qy 8576 CGCGGGAACACAGGAGATCGGGGCTTACGAGCCTTACGAGGCTATGACTAGGTA 8635
Db 6960 CGCGGGAACACAGGAGACAGGAGCCTTACGAGGCTTACGAGGCTATGACTAGTA 7019
Qy 8636 TTGCGCCCCCGGGGATCGCCCCAACCAACAGATACGACTGGAGCTGATAACATCATG 8695
Db 7020 CTCTGCCCCCTGGGACCGCCCAACCAAGATACGACTTGGAGTTGATAACATCATG 7079
Qy 8696 TTCTTCAATGTGTCAGTCGGGACAGATGATCTGGCAAAAGGTTATACTACCTACCCG 8755
Db 7080 CTCTTCAATGTGTCAGTCGGGACAGATGATCTGGCAAAAGGTTATACTATCTCACCCG 7139
Qy 8756 TGACCCACACCCCTTTCAGCGGCTGGTGGGAGACAGCTAGACACACTCCCAATCAA 8815
Db 7140 TGACCCACACCCCTTTCAGCGGCTGGTGGGAGACAGCTAGACACACTCCAGTCAA 7199
Qy 8816 CTCTTGCTAGGCAATATCATATGATGCGCCACACCTATGGGCAAGGATGATCTGAT 8875
Db 7200 TTCTTGCTAGGCAATATCATATGATGCGCCACACCTTGGGCAAGGATGATCTGAT 7259
Qy 8876 GACTCAGCTTTTCTCCATCTCTAGCTCAAGACAACTTGAAAAAGCCCTGGATTGTCA 8935
Db 7260 GACTCAGTTTCTCTCCATCTCTAGCTCAAGAACAACTTGAAAAAGCCCTTAGATTGTCA 7319
Qy 8936 GATCTAGGGGCTTGCTACTCCATAGAGCAGCTTGAACCTACCTCAGATCATTTAGCAGCT 8995
Db 7320 GATCTAGGGGCTTGCTACTCCATAGAGCAGCTTGAACCTACCTCAGATCATTTAGCAGCT 7379
Qy 8996 CCATGGCTTAGCGCAATTAACCTCAAGTACTCTCCAGTGAATCAATAGGTTGGC 9055
Db 7380 CCATGGCTTAGCGCAATTTCACTCCATAGTACTCTCCAGTGAATCAATAGGTTGGC 7439
Qy 9056 TTCAATGCTCAGGAAACTTGGGGTACCACTTTCGAACTTGGAGACATCGGGGCGAGAA 9115
Db 7440 TTCAATGCTCAGGAAACTTGGGGTACCGCTTTCGAGTCTTGGAGACATCGGGGCGAGAA 7499
Qy 9116 TGTCCGCGCTAAGCTACTGTCCAGGGGGGAGGCGGCCACTTGTGGCAGATACCTTT 9175
Db 7500 TGTCCGCGCTAAGCTACTGTCCAGGGGGGAGGCGGCCACTTGTGGCAGATACCTTT 7559
Qy 9176 TAACTGGGCACTAAGGACCAAGCTTAACTCACTCAATCCCGCGCGCTCCAGCTGGA 9235
Db 7560 CAACTGGGCACTAAGGACCAAGCTTAACTCACTCAATCCCGCGCGCTCCAGCTGGA 7619
Qy 9236 CTGTCTGGCTGGTTCGTGCTGGTTACAGCGGGGAGACATATATACAGCCTGTCTCG 9295
```

```
Db 7620 TTTATCCAGCTGGTTCTGTTGGTTTACAGCGGGGAGACATATATCAGACCTGTCTCG 7679
Qy 9296 TGCAGGACCCCGCTGGTTCCGTTGGCTTACTCTACTCTTCTGTAGGGGTAGGCATTTA 9355
Db 7680 TGCAGGACCCCGCTGGTTTCAATGTTGGTGCCTACTCTCTACTTCTGTAGGGGTAGGCATCTA 7739
Qy 9356 CTGTCTCCCAACCCAGTGAACCGGGAGCTAAACCACTCCAGGCTTAAAGCC----- 9405
Db 7740 TCTACTCCCAACCCAGTGAACCGGGAGCTTAAACACTCCAGGCCAATAGGCCATCTCTTT 7799
Qy 9406 ATTTCCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCT 9465
Db 7800 TTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCTTT 7859
Qy 9466 TCTTTTTTCTCTTCTTTTTTCCCTCTTTAAATGTTGGCTCCATCTTAGCCCTAGTCACGG 9525
Db 7860 TTTTTTCTCTTTTTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTCCTTT 7919
Qy 9526 CTAGCTGTGAAGGTCGCTGAGCGCATGACTCAGAGAGTGTGATACTTGGGCTCTCTCTG 9585
Db 7920 CTAGCTGTGAAGGTCGCTGAGCGCTTGACTGAGAGAGTGTGATACTTGGGCTCTCTCTG 7979
Qy 9586 CAGATCATGT 9595
Db 7980 CAGATCAAGT 7989

RESULT 13
US-10-509-921-14
; Sequence 14, Application US/10509921
; Publication No. US20050250093A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Hepatitis C Virus Sub-Genomic Replicons
; FILE REFERENCES: P51335
; CURRENT APPLICATION NUMBER: US/10/509,921
; CURRENT FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: 60/369,685
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 7989
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: The polynucleotide sequence encodes sequences from
; OTHER INFORMATION: HCV H77 (pBB7-SN) Replicons
US-10-509-921-14

Query Match 55.4%; Score 5312; DB 6; Length 7989;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 5650; Conservative 0; Mismatches 530; Indels 10; Gaps 1;

Qy 3416 CTTTGGCCCATCAGCGCTTCTCCCAACAAACCGGGGGTACTTGGTTCATCATCAC 3475
Db 1800 CATGGCCCTATTACGCGCTCTCCCAACAGACGCGAGGCGCTACTTGGTTCATCATCAC 1859
Qy 3476 TAGCCTCAGAGCCGCGGACAAAGACAGGTGCAAGGGGAGGTTCAGTGGTTTCTACCGC 3535
Db 1860 TAGCCTCAGAGCCGCGGACAAAGACAGGTGCAAGGGGAGGTTCAGTGGTCTCCACGC 1919
Qy 3536 AACCAATCTTCTCGGCGACCTGCATCAACGGCGTGTGTGGACTGTCTACCATGGCGC 3595
Db 1920 AACCAATCTTCTCGGCGACCTGCATCAACGGCGTGTGTGGACTGTCTATCATGTGTC 1979
Qy 3596 TGGCTCGAAGACCTTAGCGCGTCCAAAAGTCCAAATCACCCAAATGTACCAATGTAGA 3655
Db 1980 CGGCTCAAGACCTTTCGCGGCCCAAGAGGCCCAATCACCCAAATGTACCAATGTAGA 2039
Qy 3656 CTTGAGCTCGTGGCTGGCAGGCGCCCGCGGGCGCTTCCATGACACCATGAGCTG 3715
Db 2040 CCAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGTTCCCTTGACCAATGACACCTG 2099
```

Qy 3716 TGGCAGCTCGGACCTTTACTTTGGTCAACGAGACATGCTGATGTCAATTCGGGTGCGCCGGCG 3775
Db |||||
Qy 2100 CGCAGCTCGGACCTTTACTTTGGTCAACGAGCATGCCAGTCAATTCGGTTCGGCCGGCG 2159
Db |||||
Qy 3776 AGGCGACAGCGGGAAGTCTACTCTCCCGCAGGCCCTGCTCTCTACCTGAAAGGCTCTCC 3835
Db |||||
Qy 2160 GGGCGACAGCGGGAAGCTACTCTCCCGCAGGCCCTGCTCTCTACTTTGAAGGGCTCTTC 2219
Db |||||
Qy 3836 GGGTGGTCCATTTGCTTCCCTTCGGGCGAGTGTGGGGCGTCTTCGGGGTGTCTGTGTG 3895
Db |||||
Qy 2220 GGGCGGTCCATGCTCTGCCCTTCGGGCGACGCTGTGGGCATCTTTTCGGGCTGCGGTGTG 2279
Db |||||
Qy 3896 CACCCGGGGGTGCGGAAGGGGTGGACTTTCATACCCCGTTCAGTCTATGGAATCTACCAT 3955
Db |||||
Qy 2280 CACCCGAGGGGTGCGAAGGGGTGGACTTTGTACCCCGTTCAGTCTATGGAACCATCTAT 2339
Db |||||
Qy 3956 GCGGTCTCCGGTCTTTCACAGACAACTCAACCCCCCGGCTGTACCGCAGACATTTCAAAGT 4015
Db |||||
Qy 2340 GCGGTCCCGGTCTTTCACGGAACCTCGTCCCTTCGGCGGTACCGCAGACATTCAGGT 2399
Db |||||
Qy 4016 GGCACATCTGACGCTCTCTACTGGCAGCGGCNAGAGCACAAAGTGGCCGCTGCGTATGC 4075
Db |||||
Qy 2400 GGGCCATCTACACGCCCCCTACTGTGTAGCGGCAAGAGCACTAAGGTGCGCGGCTGCGTATGC 2459
Db |||||
Qy 4076 AGCCCAAGGGTACAAGGTGCTCGTCTCTGAACCCGTTCGGTTCGGCGCACCTTAGGGTTTGG 4135
Db |||||
Qy 2460 AGCCCAAGGGTATAAGGTGCTTGTCTGAACCCGTTCGGTTCGGCGCACCTTAGGGTTTGG 2519
Db |||||
Qy 4136 GCGGTATATGTCCAAAGGCACACGGTATCGACCTTAACATCAGAACTGGGGTAAAGACCAT 4195
Db |||||
Qy 2520 GCGGTATATGTCTAAGGCACATGGTATCGACCTTAACATCAGAACTGGGGTAAAGACCAT 2579
Db |||||
Qy 4196 TACGACGGCGGCTCAATTAAGTACTCACTATGGCAAGTTCCTTGGCGGACGGTGGCTG 4255
Db |||||
Qy 2580 CACCAAGGGTGGCCCATCACTACTCTCACTATGGCAAGTTCCTTGGCGGACGGTGGTGG 2639
Db |||||
Qy 4256 TTCTGGGGCGGCTATGACATCAATAATATGTGATGATGCCACTCACTCACTGCTGCATAC 4315
Db |||||
Qy 2640 CTCTGGGGCGGCTATGACATCAATAATATGTGATGATGCCACTCACTCACTGCTGCATAC 2699
Db |||||
Qy 4316 CATCTTTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGACGGTGGAGCGCGGCTCGTGT 4375
Db |||||
Qy 2700 TATCTTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGCGGCTGGAGCGGACTCGTGT 2759
Db |||||
Qy 4376 GCTCGGCACCGCTACACTCTCGGGATCGGTTACGGTGCCACACCCCAATATCGAGGAAT 4435
Db |||||
Qy 2760 GCTCGGCACCGCTACCGGCTCTCGGGATCGGTCAACCGTGCCACATCCAAACATCGAGGAGGT 2819
Db |||||
Qy 4436 AGGCTGTCCAAATGAGAGATCCCTTCTATGCGCAAGGCCATCCCATTTGAGGCCAT 4495
Db |||||
Qy 2820 GGTCTGTCCAGCATGGAAGAAATCCCCCTTTATGGCAAGGCCATCCCCATCGAGACCAT 2879
Db |||||
Qy 4496 CAAAGGGGGGAGGCATCTCATTTTCTGCGCATTTCCAAAGAAATGTGACGAGCTCGCGC 4555
Db |||||
Qy 2880 CAAAGGGGGGAGGCACCTCATTTTCTGCGCATTTCCAAAGAAATGTGATGAGTCTCGCGC 2939
Db |||||
Qy 4556 AAAGTGTACAGGCCTCGGACTGAAACGCTGTAGCATATTTACCGGGGCTTGATGTGTCGT 4615
Db |||||
Qy 2940 GAAAGTGTCCGGCTCGGACTCAATGCTGTAGCATATTTACCGGGGCTTTGATGTATCGGT 2999
Db |||||
Qy 4616 CATACGGCTATCGGAGAGTGTGTGTGGGCAACAGACGCTCTAATGACGGGTTTTCAC 4675
Db |||||
Qy 3000 CATACAACATAGCGGAGAGCTCATTTGTGTGACAAACGAGACGCTCTAATGACGGGCTTTTAC 3059
Db |||||
Qy 4676 CGGCGATTTTGAATCAGTCACTGCAATATGATGTGTACCCAGACAGTTCGACTTCAG 4735
Db |||||
Qy 3060 CGGCGATTTTGAATCAGTCACTGCAATATGATGTGTACCCAGACAGTTCGACTTCAG 3119
Db |||||
Qy 4736 CTTTGAATCCCATCTTCAACATTTGAGACGACCGGTGCCCAAGACGCGGTGTGCGGCTC 4795
Db |||||
Qy 3120 CTTGACCCGACCTTCAACATTTGAGACGACCGGTGCCCAAGACGCGGTGTGACGCTC 3179
Db |||||

Qy 4796 GCAACGCGAGGTTAGAACTGGCAGGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGG 4855
Db |||||
Qy 3180 GCAGCGCGAGGACGAGTGTAGGGGACGAGTGGGATTTACAGGTTTGTGACTCCAGG 3239
Db |||||
Qy 4856 AGAACGGCCCTCGGGCATGTTCCGATTTCTTGGTCTCTGTGTGAGTGTCTATGACGGGGCTG 4915
Db |||||
Qy 3240 AGAACGGCCCTCGGGCATGTTCCGATTTCTTGGTCTCTGTGTGAGTGTCTATGACGGGGCTG 3299
Db |||||
Qy 4916 TGTCTGTATGAGTCTACGCGCGCTGAGACCTCGGTTAGGTTGGGGCTTACCTAAATATAC 4975
Db |||||
Qy 3300 TGTCTGTATGAGTCTACGCGCGCTGAGACCTCTAGTTAGGTTGGGGCTTACCTAAATATAC 3359
Db |||||
Qy 4976 ACCAGGGTTGCGGCTCTGCGCAGGACCATCTGGAGTCTTGGGAGAGCGCTTTCACAGGCT 5035
Db |||||
Qy 3360 ACCAGGGTTGCGGCTCTGCGCAGGACCATCTGGAGTCTTGGGAGAGCGCTTTCACAGGCT 3419
Db |||||
Qy 5036 CACCCATATGATGCCCATCTCTCTGTCCTCCAGACTAAACAGCAGGAGACAACTTTTCTTTA 5095
Db |||||
Qy 3420 CACCCATATGAGCGCCCATTTCTTTGTCCTCAGACTTAAGCAGCAGGAGACAACTTTCCCTA 3479
Db |||||
Qy 5096 CTTGTTGSCATATCAAGCTACAGTGTGCGCAGGCTCAAGCTCCACTCCATCTGTTGGGA 5155
Db |||||
Qy 3480 CTTGTTGAGCATACAGGCTACGCTGTCGCGAGGCTCAGGCTCCACTCCATCTGTTGGGA 3539
Db |||||
Qy 5156 CCAAATGTGGAAGTGTCTCATACGCTGAAAACCTTACACTGACGCGGCCAAACACCCCTGCT 5215
Db |||||
Qy 3540 CCAAATGTGGAAGTGTCTCATACGCTTAAAGCTTACGCTGACGCGGCCAAACACCCCTGCT 3599
Db |||||
Qy 5216 GTATAGGCTAGGAGCCGTCCAAAATAGGTCATCTCTCACACACCCCATTAATAATATAC 5275
Db |||||
Qy 3600 GTATAGGCTAGGAGCCGTTCAAAACGAGGTTACTTACACACACACCCCATTAATAATATAC 3659
Db |||||
Qy 5276 CATGGCATGATGTCGGCTGACCTGAGGTGCTCACTAGCACCTGGGTGCTGTGAGGGG 5335
Db |||||
Qy 3660 CATGGCATGATGTCGGCTGACCTGAGGTGCTCACTAGCACCTGGGTGCTGTGAGGGG 3719
Db |||||
Qy 5336 AGTCTTTGCGAGCTTTGGCCGCATACTGCTGACACAGCAGGAGTGTGCTCATTTGTTGGGCGAG 5395
Db |||||
Qy 3720 AGTCTTAGCAGCTCTGCGCGCGTATTGCTTGACAAACAGGACGCTGGTTCATTTGTTGGGCGAG 3779
Db |||||
Qy 5396 GATCATCTTGTTCGGGAAGCAGCTGCTGTTTCCGACAGGGAAGTCTCTTACAGGAGTGT 5455
Db |||||
Qy 3780 GATCATCTTGTTCGGGAAGCAGGCGCATCATTTCCGACAGGGAAGTCTCTTTTACCGGAGTGT 3839
Db |||||
Qy 5456 CGATGAGTGAAGAGTGTGCTCACAACCTTCTTACATCGACGAGGAAGTGCAGTCTGC 5515
Db |||||
Qy 3840 CGATGAGTGAAGAGTGTGCGCTCACACCTCTCTTACATCGAACAGGGAATGCAGTCTGC 3899
Db |||||
Qy 5516 CGAGCAATTCAGACAAAGGCGCTCGGGTGTGTCAAACGCGCCACCAAGCAAGCGGAGGC 5575
Db |||||
Qy 3900 CGACAAATTCAAACAGAAAGCAATCGGGTGTCTGCAACAGCCACCAAGCAAGCGGAGGC 3959
Db |||||
Qy 5576 TGTCTCTCCGTGTGAGTCCAAAGTGGCGAGCCCTTTGAGACCTTCTGGGCGAAGCAAT 5635
Db |||||
Qy 3960 TGTCTCTCCGTGTGGAATCCAAAGTGGCGGACCTCTGAAGCTTCTGCGGCGAAGCATAT 4019
Db |||||
Qy 5636 GTGGAATTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGGAACCC 5695
Db |||||
Qy 4020 GTGGAATTCATCAGCGGATACAAATTTTACAGGCTTGTCCACTCTGCTGCGCAAGCC 4079
Db |||||
Qy 5696 CGCGATAGCATCATTTGATGGCATTTTACAGCTTCTATCACTAGCCGCTCACACCCAAAA 5755
Db |||||
Qy 4080 CGCGATAGCATCATTTGATGGCATTTTACAGCTTCTATCACTAGCCGCTCACACCCAAACA 4139
Db |||||
Qy 5756 CACCTCTCTGTTTAACTCTTTGGGGGATGGGTGGTCCCAACTCGCTCTCTCCAGGCG 5815
Db |||||
Qy 4140 TACCTCTCTGTTTAACTCTCTGGGGGATGGGTGGCGCCGCCCAACTTGTCTCTCCAGGCG 4199
Db |||||
Qy 5816 TGGCTCAGCTTTCTGGGGCGGCGCATCGCGGAGCGGCTTGTGGCAGCATAGGCTTTGG 5875
Db |||||
Qy 4200 TGTCTCTCTTCTGTAGCGCGCGCATCGCTGGAGCGGCTGTGTGGCAGCATAGGCTTTGG 4259
Db |||||
Qy 5876 GAAAGTGTCTGTGGACATCTTTGGCGGGCTATTTGGGGCAGGGGTAGCGCGGCACACTCTGTGC 5935
Db |||||

4260	Db	 GAAGGTGCTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGCGCTCGTGCGC
5936	Qy	 CUTTAAAGGTCATGAGCGGGCGAGTGCCTCCACCGAGGACCTGGTCAACTTACTTCCCTGCG
4320	Db	 CUTTAAAGGTCATGAGCGGGCGAGTGCCTCCACCGAGGACCTGGTTAACTTACTTCCCTGCG
5996	Qy	 CATCTCTCTCTCTGGTGCCTCGTCTCGGGGTCTGTGTCGCGAGCAATACTCGTCTCGCA
4380	Db	 TATCTCTCTCCCTGGCGCCTAGTCTGTGGGGTCTGTGTGCGCAGCGATACTGCGTCTCGCA
6056	Qy	 CGTGGGCGCGGAGAGAGGGGGCTGTGCAGTGGATGAACCGGCTGTAGCGTTGCGTTCGCG
4440	Db	 CGTGGGCGCGGAGAGGGGGCTGTGCAGTGGATGAACCGGCTGTAGCGTTGCGTTCGCG
6116	Qy	 GGGTAAACACGCTCTCCCTTACGCACCTATGTGCTGAGAGCGAGCGTTCAGCAGCGTGTCA
4500	Db	 GGGTAAACACGCTCTCCCGCAGCAGCTATGTGCTGAGAGCGAGCTTCAGCAGCGTGTCA
6176	Qy	 TCAGATCTCTCTAGGCTTTACCATCACTCAACTGCTGAAGCGGCTCCACAGTGGATTAA
4560	Db	 TCAGATCTCTCTAGTCTTTACCATCACTCAGCTGCTGAAGAGCTTCCACAGTGGATCAA
6236	Qy	 TGAGGACTGCTCTPACGCCATGCTTCCGGCTCTGGCTTAAGGATGTTTGGGATTTGGATATG
4620	Db	 CGAGGACTGCTCCACGCCATGCTCCGGCTCTGGCTTAAGAGATGTTTGGGATTTGGATATG
6296	Qy	 CAGGCTGTTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCTCTGCCCGGTTACCGGG
4680	Db	 CAGGCTGTTGACTGATTTTCAAGACCTGGCTCCAGTCCAACTCTCTGCCCGGTTACCGGG
6356	Qy	 AGTCCCTTTCTCTCATGCCAACCGGGTCAAGGGAGTCTTGCGGGGGGAGCGCATCAT
4740	Db	 AGTCCCTTTCTCTCATGTCAAAGTGGGTACAGGGAGTCTGGCGGGCGAGGGCATCAT
6416	Qy	 GCAAAACACCTGCGCCATGCGGAGCAGATCGCGGACATGTCAAAAACGGTTCCATGAG
4800	Db	 GCAAAACACCTGCGCCATGTTGGAGCAGATCAGCGGACATGTAAAAACGGTTCCATGAG
6476	Qy	 GATCGTAGGCGCTAGAACCTGACAGCAACAGTGGCAGCGGAACGTTCCCATCAACGCATA
4860	Db	 GATCGTAGGCGCTAGAACCTGTAGTAAACAGTGGCATGGAACTTCCCATTAACCGTGA
6536	Qy	 CACACGGGACCTTGACACCCCTCCCGCGCGCCAACTATTTCAGGGCGCTATGGCGGGT
4920	Db	 CACACGGGCGCTGACGCGCCTCCCGCGCGCCAAATATTCTTAGGGCGCTGTGGCGGGT
6596	Qy	 GGCTGCTGAGGAGTACGTGAGGTTACGCTGTGGGGATTTCCACTGATGACGGGCAT
4980	Db	 GGCTGCTGAGGAGTACGTGAGGTTACGCGGGTGGGGATTTTCCACTGATGACGGGCAT
6656	Qy	 GACCACTGACAAAGTAAAGTGCACAGTCCAGGTTCCGGCCCCCGAATTTCTTACGGAGGT
5040	Db	 GACCACTGACAAAGTAAAGTGCAGGTTCCGGGGTGGGGATTTTCCACTGATGACGAGGT
6716	Qy	 GGATGGAGTGGGTTGCACAGGTACGCTCCGGGTGCAAACTCTTCTACGGGAGGACGT
5100	Db	 GGATGGGTTGGGTTGCACAGGTACGCTCCAGGTGCAAACTCTTCTACGGGAGGAGGT
6776	Qy	 CACGTTCCAGTCCGGGTCAAACCAATATCTTGGTGGGTCCAGCTCCCATGCGAGGCCGA
5160	Db	 CACATCTCGTGGGTCATCAATCAATACCTGGTTGGGTCCAGCTCCCATGCGAGGCCGA
6836	Qy	 ACCGGACGTAAAGTGTCTTACTTCATGCTCACCGATCCCTCCCAATTAACAGCAGAC
5220	Db	 ACCGGACGTAGCAGTGTCTCACTTCATGCTCACCGACCCCTCCCAATTAACGGCGGAGC
6896	Qy	 GGCTAAGCGTAGGCTGAGAGGCTCTCCCGCTCTTTAGCCAGCTCATCAGCTAGCCA
5280	Db	 GGCTAAGCGTAGGCTGAGAGGATCTCCCGCTCTTGGCAGCTCATCAGCTATCCA
6956	Qy	 GTTGCTGCGCTCTTTTGAAGGCGCATGCACTACCACTGACCTCCCGGACGCTGA
7015	Qy	

5340	GTGTCTGCGCCTTCTCTTGAAGGCAACATGCACACTACCGGTATGACTCCCCGGACGCTGA	5399
7016	CCTCATCGAGGCAACCTCTTTGTGGCGGAGGAGATGGSCGGAACAATCATCTCGCGTGA	7075
5400	CCTCATCGAGGCAACCTCTCTGTGGCGGAGGAGATGGGCGGAACATCACCCGCGTGA	5459
7076	GTGAGAAATAGGTAGTAAATTTCTGGACCTCTTTGGAACCGCTTTCACGGAGGGGATGA	7135
5460	GTGAGAAATAGGTAGTAAATTTTGAGACTCTTTTCGAGCCGCTTCCAAGCGGAGAGATGA	5519
7136	GAGGAGATATCCGTTCGCGCGGAGATCTCTGCGAAATCCAGGAAGTTTCCCTCAGCGTT	7195
5520	GAGGGAAGTATCCGTTCGCGCGGAGATCTTCGCGAGGTCAGGAATTCCTCGAGCAT	5579
7196	GCCCATATGGCAAGCCCGGACTTACAATCTCTCACTGTCTAGAGTCTCTGGAAGAACCGGA	7255
5580	GCCCATATGGCAAGCCCGGATTAACAACCTCCACTGTTAGAGTCTCTGGAAGAACCGGA	5639
7256	CTAGTCTCCCTCGGTGGTACACGGATGCCATTTGCCACCTACCAAGGCTCTTCCAATACC	7315
5640	CTAGTCTCCCTCAGTGGTACACGGGTGTCTCAATTTCCGCGCTTCCAAGGCCCTCCGATACC	5699
7316	ACCTCCACGAGAAAGAGACGGTGTCTCTGACAGAAATCCAATGTGTCTTCTGCCTTGGC	7375
5700	ACCTCCACGAGAAAGAGACGGTGTCTCTGTGACAAATCTACCGTGTCTTCTGCCTTGGC	5759
7376	GGAGCTCGCCACTAAGACCTTTGGTAGTCTCCGGATCGTCCGGCGGTTGATAGCGGACGGC	7435
5760	GGAGCTCGCCACAAGACCTTCGGCAGCTCCGAATCGTCCGGCGCTCGACAGCGCACGGC	5819
7436	GACGGCCCTTCGTGACCTGGCCTCCAGACGAGTGAACAAGGATCCGAGCTTGAGTCTGA	7495
5820	AACGGCCCTCTCTGACACAGCCCTCCAGACGCGGAGCCGGGATCCGAGCTTGAGTCTGA	5879
7496	CTCTCCATCCGCCCTTTGAAGGGAGCGCGGGGACCCCGATCTCAGGCGACGGGTCTTG	7555
5880	CTCTCATGCCCTCTTTGAGGGGAGCGGGGGATCCCGATCTCAGCGACGGGTCTTG	5939
7556	GTCTACCGTAGTAGAGGCTAGTGAGGATGTCTGTCTGCTCAATGTCTCTATACGTG	7615
5940	GTCTACCGTAAGCGAGGAGCTAGTGAGGACGTCTGTCTGTCTGATGTCTCTACACATG	5999
7616	GACAGGGCCCTGATCACGCCATTCGCTGCGGAGGAAGTTAAGCTCCCATCAACCGCTT	7675
6000	GACAGGGCCCTGATCACGCCATTCGCTGCGGAGGAACCAAGCTGCCATCAATGCATCT	6059
7676	GAGCAACTCTTTGCTGCTCACCAACATGTGTCTACGCCACAACATCCCGCAGCGCAAG	7735
6060	GAGCAACTCTTTGCTCGTCCACCAACTTGTGTCTATGCTATACAACTCTCGCAGCGCAAG	6119
7736	CCTCCGGCAGAAGGTCACTTTGACAGATTTGCAAGTCTCGATGATCATTTACCGGGA	7795
6120	CCTCCGGCAGAAGGTCACTTTTGACAGACTGCAGTCTCGAGTCCGACCACTACCGGGA	6179
7796	CGTACTCAAGGAGATGAAGCGAAGCGGTCCACAGTTAAGGCTTAAGCTTCTATCTATAGA	7855
6180	CGTCTCAAGGAGATGAAGCGAAGCGGTCCACAGTTAAGGCTTAAGCTTCTATCTCGTGA	6239
7856	GGAGGCTCGAAGCTGACGCCCCACATTCGGGCAAAATCCAAATTTGGCTTATGGGGCAAA	7915
6240	GGAGGCTCGAAGCTTATCCAGCAAGCGGTTCACCAATTCGGCCAGATCTAATTTGGGCAAA	6299
7916	GGAGCTCCGGAACCTATCCAGCAGGGCCGTTAAACCATCATCCGTCCGTGTGGGAGACTT	7975
6300	GGAGCTCCGGAACCTATCCAGCAAGCGGTTCACCAATTCGGCCAGATCTAATTTGGGCAAA	6359
7976	GCTGGAGACACTGAACAACAATTTGACCAACCATCATCGCAAAAGTGTAGGTTTCTG	8035
6360	GCTGGAGACACTGAGACACCAATTTGACCAACCATCATGGCAAAATGAGGTTTTCTG	6419
8036	CGTCCACCAAGAGAGGAGCGCCGAAGCCAGCTCGCTTATCGTATTTCCAGACCTCGG	8095
6420	CGTCCACCAAGAGAGGAGCGCCGAAGCAGCTCGCTTATCGTATTTCCAGATTTGGG	6479

QY	3596	TGCTCGAAGACCTTACCGCGTCCAAAGGTCCAATCACCCAAATGTACACCAATGTAGA	3655
Db	1980	CGGCTCAAGACCTTTCGCGGCCAAAGGGCCCAATCACCCAAATGTACACCAATGTAGA	2039
QY	3656	CTTGAGACTTGTGGCTGGCAGGCGCCCGCGGGCGGCTTCATCAGACCAATGTACAGCTG	3715
Db	2040	CCAGGACCTTGTGGCTGGCAGGCGCCCGCGGGCGGCTTCATCAGACCAATGTACAGCTG	2099
QY	3716	TGGCAGCTCGGACCTTACTTGTGTACAGACATGTGTATCTTCCGGTGGCGCGCG	3775
Db	2100	CGGACCTCGGACCTTACTTGTGTACAGACATGTGTATCTTCCGGTGGCGCGCG	2159
QY	3776	AGCGCAGCAGGAGGAGTCTACTTCCCGCAGCGCGCTCTCTACTGAAAGGCTCTTC	3835
Db	2160	GGCGCAGCAGGAGGAGGCTACTTCCCGCAGCGCGCTCTCTACTGAAAGGCTCTTC	2219
QY	3836	GGGTGTCTCATTTGTCCTTGGCGGACAGCTGTGGGGCTCTTCCGGCTGCTGTG	3895
Db	2220	GGCGGTCTCATTTGTCCTTGGCGGACAGCTGTGGGGCTCTTCCGGCTGCTGTG	2279
QY	3896	CACCGGGGGTTCGGAAGCGGTGGACTTATACCGGTTGAGTCTATGGAATACATCAT	3955
Db	2280	CACCGGGGGTTCGGAAGCGGTGGACTTATACCGGTTGAGTCTATGGAATACATCAT	2339
QY	3956	GCCTGTCTCCGGTCTTTCACAGACAACTCAACCCCGCGGCTGTACCGCAGACATTTCAAGT	4015
Db	2340	GCCTGTCTCCGGTCTTTCACAGACAACTCAACCCCGCGGCTGTACCGCAGACATTTCAAGT	2399
QY	4016	GGCAGATCTGACGCTCTACTGCGAGCGGAGAGCACAAGTGCAGGCTGCGGTATGC	4075
Db	2400	GGCCCATCTACAGCCCTTACTGTAGCGGCAAGAGACATAAGGTGCGGCTGCGGTATGC	2459
QY	4076	AGCCAGGCTACAGGCTGCTCGTGAACCGTTCGTCGCGCACCTTAGGTTTGG	4135
Db	2460	AGCCAGGCTACAGGCTGCTCGTGAACCGTTCGTCGCGCACCTTAGGTTTGG	2519
QY	4136	GGGTATATGTCCAAAGGCACACGGTATCGACCTTAACATCAGAACTGGGGTAAAGCAAT	4195
Db	2520	GGGTATATGTCCAAAGGCACACGGTATCGACCTTAACATCAGAACTGGGGTAAAGCAAT	2579
QY	4196	TACACCGGGGCTTCAATAGTACTCACTATATGGCAAGTTCCTTGGCGAGCGTGCTG	4255
Db	2580	CACACGGGTGCCCCATCAGTACTCACCTATATGGCAAGTTCCTTGGCGAGCGTGCTG	2639
QY	4256	TTCTGGGGGCGCTTATGACATCAATATGTATGTAGTGCACCTCAACTGACTCGACTAC	4315
Db	2640	CTCTGGGGGCGCTTATGACATCAATATGTATGTAGTGCACCTCAACTGACTCGACTAC	2699
QY	4316	CATCTTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGACGGCTGGAGCGCGCTCGTCTG	4375
Db	2700	TATCTTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGACGGCTGGAGCGCGCTCGTCTG	2759
QY	4376	GCTCGCCACCGCTACACTCCGGGATCGGTTACCGTGGCCACACCCCAATATCGAGGAAT	4435
Db	2760	GCTCGCCACCGCTACACTCCGGGATCGGTTACCGTGGCCACACCCCAATATCGAGGAGT	2819
QY	4436	AGGCTGTCCAAATATGAGAGATCCCTTCTATGGCAAGCCATCCCATTTAGGCGCAT	4495
Db	2820	GGCTGTGTCCAGCACTGGAGAAATCCCTTCTATGGCAAGCCATCCCATTTAGGCGCAT	2879
QY	4496	CAAGGGGGAGGACCTCATTTTCTGCGCATTCGAAGAGAAATGTGACGAGCTCGCGC	4555
Db	2880	CAAGGGGGAGGAGGACCTCATTTTCTGCGCATTCGAAGAGAAATGTGATGAGCTCGCGC	2939
QY	4556	AAAGCTGACAGGCTTCGAGCTGAACGCTGTAGCATATTTACCGGGGCTTGTATGTCCGT	4615
Db	2940	GAAGCTGTCCGGCTTCGAGCTCAATGTGTAGCATATTTACCGGGGCTTGTATGTCCGT	2999
QY	4616	CATACCGGCTTATCGGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4675
Db	3000	CATACCAACTAGCGGAGACGCTCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3059

QY	4676	CGGCGATTTTGACTCAGTGATCGACTGCAATACATGTGTACCCAGACAGTTCGACTTCAG	4735
Db	3060	CGGCGATTTTGACTCAGTGATCGACTGCAATACATGTGTACCCAGACAGTTCGACTTCAG	3119
QY	4736	CTTTGATCCCACTTACCATTGAGACGACGCGGTGCCCAAGACGCGGTGTTCGCGCTC	4795
Db	3120	CTTTGATCCCACTTACCATTGAGACGACGCGGTGCCCAAGACGCGGTGTTCGCGCTC	3179
QY	4796	GCAA CGCGAGGTAGAACTGGCAGGGTAGGATGSCATCTACAGCTTTGTGACTCCAGG	4855
Db	3180	GCAGCGCGAGGACGAGCTGGTAGGGCAGATGGGCATTTACAGCTTTGTGACTCCAGG	3239
QY	4856	AGAA CGCGGCTCGGGCATGTTTGATTTCTCGGTCTGTGTGAGTGCTATGACCGCGGCTG	4915
Db	3240	AGAA CGCGGCTCGGGCATGTTTGATTTCTCGGTCTGTGTGAGTGCTATGACCGCGGCTG	3299
QY	4916	TGCTTGTATGAGCTCACGCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACCTAAATAC	4975
Db	3300	TGCTTGTATGAGCTCACGCCGCTGAGACCTCAGTTAGGTTGCGGGCTTACCTAAATAC	3359
QY	4976	ACCAGGTTGCGGCTGTGCGCAGACCATCTGGAGTTCTGGGAGAGCTTTCACAGGCT	5035
Db	3360	ACCAGGTTGCGGCTGTGCGCAGACCATCTGGAGTTCTGGGAGAGCTTTCACAGGCT	3419
QY	5036	CACCCACATAGATCCCACTTCTGTCGCCAGACTAAACAGGACGAGAGACAACTTTCCTTA	5095
Db	3420	CACCCACATAGATCCCACTTCTGTCGCCAGACTAAACAGGACGAGAGACAACTTTCCTTA	3479
QY	5096	CTGTGTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCCATCGTGGGA	5155
Db	3480	CTGTGTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCCATCGTGGGA	3539
QY	5156	CCAATGTGGAAGTGTCTCATCGGCTGAAACCTACACTGCA CGGCGCAACACCCCTGCT	5215
Db	3540	CCAATGTGGAAGTGTCTCATCGGCTGAAACCTACACTGCA CGGCGCAACACCCCTGCT	3599
QY	5216	GTATAGCTAGGACCGTCCAAATGAGGTATCTCACACCCCACTCACTAAATACAT	5275
Db	3600	GTATAGCTAGGACCGTCCAAATGAGGTATCTCACACCCCACTCACTAAATACAT	3659
QY	5276	CATGGCATGATGCTCGGCTGACCTGGAGTCTGCTACCTAGCACCTGGGTGCTGTAGCGG	5335
Db	3660	CATGGCATGATGCTCGGCTGACCTGGAGTCTGCTACCGAGCACCTGGGTGCTGTAGCGG	3719
QY	5336	AGTCTTGTGAGCTTTGGCGCATATGCTCTGACGACAGGAGTGTGTCTATGTGGGAG	5395
Db	3720	AGTCTTGTGAGCTTTGGCGCATATGCTCTGACGACAGGAGTGTGTCTATGTGGGAG	3779
QY	5396	GATCATCTTGTCCGGGAGCGAGCTGCTGTTCCGACAGGAGTCTCTTACCAAGGAGTT	5455
Db	3780	GATCATCTTGTCCGGGAGCGAGCTGCTGTTCCGACAGGAGTCTCTTACCAAGGAGTT	3839
QY	5456	CGATGAGATGGAAGTGTGCTCACAACTTCTTACATCGAGCAGGGAATGCACTCGC	5515
Db	3840	CGATGAGATGGAAGTGTGCTCACAACTTCTTACATCGAGCAGGGAATGCACTCGC	3899
QY	5516	CGAGCAATTTCAAGCAAAAGGCTCGGGTTGTTGCAAAACCGGCCAACGAAGCGGAGGC	5575
Db	3900	CGAGCAATTTCAAGCAAAAGGCTCGGGTTGTTGCAAAACCGGCCAACGAAGCGGAGGC	3959
QY	5576	TGCTGCTCCGCTGTGAGTCCAAAGTGGCGAGCCCTTGAGACCTTCTGGGCGAGACAT	5635
Db	3960	TGCTGCTCCGCTGTGAGTCCAAAGTGGCGAGCCCTTGAGACCTTCTGGGCGAGACAT	4019
QY	5636	GTGGAATTTCAATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTGCAACCC	5695
Db	4020	GTGGAATTTCAATCAGCGGAATACAGTATTTAGCAGGCTTGTCCACTCTGCTGCAACCC	4079
QY	5696	CGCGATAGCATCATTTGATGGCATTTACAGCTTCTTATCACTAGCCCGCTTACCAACCCAAA	5755
Db	4080	CGCGATAGCATCATTTGATGGCATTTACAGCTTCTTATCACTAGCCCGCTTACCAACCCAAA	4139
QY	5756	CACCTCTCTGTTTAAACATCTTTGGGGGAGTGGTGGCTGCCCCAACCTCGCTCTCTCCAGCGC	5815

4140	Db	 TACCTCTCTGTTTAAACATCTCGGGGAGTGGGTGGCGCCCAACTTGTCTCTCCAGCGC	4199
5816	Qy	TGCGTAGCTTTCGTGGGCGCCGGCATCGCCGAGCGGCTGTTGGCAGCATAGGCTCTGG	5875
4200	Db	TGCTTCTGCTTTCGTAGCGCCGGCATCGCTGAGCGGCTGTTGGCAGCATAGGCTCTGG	4259
5876	Qy	GAAGGTGCTCTGTGGACATCTTGGCGGGCTATGGGGCAGGGGTAGCCGGCGCACTCGTGGC	5935
4260	Db	GAAGGTGCTTGTGATATTTTGGCAGGTTATGGACAGGGGTGGCAGGCGCTCGTGGC	4319
5936	Qy	CTTTAAGGTCTATGACGGCGCAGGTGCCCTCCAACGAGGACCTGGTCAACTTACTCTCCTGC	5995
4320	Db	CTTTAAGGTCTATGACGGCGCAGATGCCCTCCACCGAGGACCTGGCTAACTACTCTCCTGC	4379
5996	Qy	CATCCTCTCTCTGGTGCCTTGGTCTGGGGTCTGTGCGCAGCAATACTGCGTGGCA	6055
4380	Db	TATCCTCTCTCCCTGGCGCCCTAGTCTGTCGGGGTCTGTGCGCAGCATCTGCGTCGGA	4439
6056	Qy	CGTGGGCCCCGGAGAGGGGGCTGTGCAGTGGATGAACCCGGCTGATAGCGTTCCCTTCGCG	6115
4440	Db	CGTGGGCCCCAGGGAGGGGGCTGTGCAGTGGATGAACCCGGCTGATAGCGTTCCCTTCGCG	4499
6116	Qy	GGGTAAACAAGTCTCCCTTAGCACTATGTGCTTAGAGCGACGCTGCAGCACTGTGTAC	6175
4500	Db	GGGTAAACAAGTCTCCCTTAGCACTATGTGCTTAGAGCGACGCTGCAGCACTGTGTAC	4559
6176	Qy	TCAGATCCTCTCTAGCTTACCATCATCTCACTGCTGAAGGGCTCCACAGTGGATTAA	6235
4560	Db	TCAGATCCTCTCTAGCTTACCATCATCTCACTGCTGAAGGGCTCCACAGTGGATTAA	4619
6236	Qy	TGAGGACTGCTCTACGCCATGCTCCGGCTCGTGGCTAAGGATGTTTGGGATTTGGATG	6295
4620	Db	CGAGGACTGCTCTACGCCATGCTCCGGCTCGTGGCTAAGGATGTTTGGGATTTGGATG	4679
6296	Qy	CACGGTGTGTACTCAAGACCTGGCTCCAGTCCAACTCTCTGCCCGGGTTACCGGG	6355
4680	Db	CACGGTGTGTACTCAAGACCTGGCTCCAGTCCAACTCTCTGCCCGGGTTACCGGG	4739
6356	Qy	AGTCCCTTCTCTCATGCGCAACCGGGGTACAGGGAGTCTGGCGGGGGACCGCATCAT	6415
4740	Db	AGTCCCTTCTCTCATGCGCAACCGGGGTACAGGGAGTCTGGCGGGGGACCGCATCAT	4799
6416	Qy	GCAAAACACCTGCCCATCGCGAGCACAGATCGCCGGACATGTCAAAACGGTTCATGAG	6475
4800	Db	GCAAAACACCTGCCCATGTGGAGCACAGATCACCGGACATGTGAAAACGGTTCATGAG	4859
6476	Qy	GATCGTAGGGCTAGAACCTGCAGCAACACGTGGCACGGAAACGTTCCCCATCAACGATA	6535
4860	Db	GATCGTAGGGCTAGAACCTGTAGTAAACACGTGGCATGGAACATTCCTCCATTAAACGGTA	4919
6536	Qy	CACACAGGACCTTGACACACCTCCCGCGGCCAACTATTTCACAGGGCGCTATGGCGGGT	6595
4920	Db	CACCAAGGGCCCTGACAGCCCTCCCGCGGCCAAATTTATTCTAGGGCGCTGTGGCGGGT	4979
6596	Qy	GGCTGCTGAGGAGTACGTGGAGGTTACCGGGGTGGGGATTTCCACTACGTGACGGGCAT	6655
4980	Db	GGCTGCTGAGGAGTACGTGGAGGTTACCGGGGTGGGGATTTCCACTACGTGACGGGCAT	5039
6656	Qy	GACCACTGACAAACGTAAAGTGCCTATGCAAGTTCGGGCCCCCGAAATTTCTTACGGAGGT	6715
5040	Db	GACCACTGACAAACGTAAAGTGCCTATGCAAGTTCGGGCCCCCGAAATTTCTTACAGAAGT	5099
6716	Qy	GGATGAGTGGGTTGACACGATACGCTCCGGGTGCAAACTCTTCTTACGGGAGGACGT	6775
5100	Db	GGATGGGTTGGTTGACACGATACGCTCCGGGTGCAAACTCTTCTTACGGGAGGACGT	5159
6776	Qy	CACGTTCCAGGTCCGGCTCAACCAATACTTGGTGGGTGCGAGCTCCCATGCGAGCCCGA	6835
5160	Db	CACATTCCTGGTCCGGCTCAATCAATAGCTGGTGGGTGCAAGCTCCCATGCGAGCCCGA	5219
6836	Qy	ACCGGACCTTAACAGTGTCTTACTTCCATGCTCACCGATCCCTTCCCAACATTAACAGCAGAC	6895

Db	5220	ACCGGACGTAGAGT	GCTC	ACTTCCAT	TGCTC	ACCGACG	CCCTCC	CAAT	TACGGCGG	AGAC	5279
Qy	6896	GGCTAAGCGTAGCT	GGCTAG	AGGGTCT	CCCCCTCT	TTT	TAGCCAG	CTCAT	CAGCTAG	CCCA	6955
Db	5280	GGCTAAGCGTAGCT	GGCCAG	GGGATCT	CCCCCTCT	TTT	GGCCAG	CTCAT	CAGCTAG	CCCA	5339
Qy	6956	GTTGTCTGCGCCTT	CTTTGA	GCGGAC	TACGC	ACTAC	CCACCAT	GATCT	CCCGGAC	CGCTGA	7015
Db	5340	GCTGTCTGCGCCTT	CTCTTGA	AGGCAAC	ATGCA	CTAC	CCCGCTCAT	GCTC	CCCGGAC	CGCTGA	5399
Qy	7016	CCTCATCGAGGCCAA	CCTCTT	TGGCGG	CAGGAG	ATGGG	CGGAAA	CATCA	CTCGCTGG	GA	7075
Db	5400	CCTCATCGAGGCCAA	CCTCTG	TGGCGG	CAGGAG	ATGGG	CGGAAA	CATCA	CTCGCTGG	GA	5459
Qy	7076	GTGAGAAATAAGGT	AGTAA	TCTTGG	ACTCTT	TCGAA	CCGCTT	CA	CGCGAG	GGGATGA	7135
Db	5460	GTGAGAAATAAGGT	AGTAA	TCTTGG	ACTCTT	TCGAG	CCGCTT	CA	CGCGAG	GGGATGA	5519
Qy	7136	GAGGAGATATCCGT	TCGCGG	CGAGAT	TCCTG	CGAAAT	TC	CAGGAA	GTTCC	CTCAGGCTT	7195
Db	5520	GAGGAAATATCCGT	TCGCGG	CGAGAT	TCCTG	CGGAGT	TC	CAGGAA	TTC	CTCGAGCGAT	5579
Qy	7196	GCCATATGGGAC	CGCCG	ACTAC	ATCTCT	CAC	TCTAG	AGTCT	CTGGA	AGGACCGCGA	7255
Db	5580	GCCATATGGGAC	CGCCG	ACTAC	ATCTCT	CAC	TCTAG	AGTCT	CTGGA	AGGACCGCGA	5639
Qy	7256	CTAGCTCCCTCCG	TGGTG	TACAC	GGATGCC	ACTT	GGCCACT	TAC	CAAGG	CTCCTCAAT	7315
Db	5640	CTAGCTCCCTCCG	TGGTG	TACAC	GGG	TGTCAT	TG	CCGCTT	GCCA	AGGCTCCGAT	5699
Qy	7316	ACCTCCACGAGAA	AAGAG	CGGTTG	CTG	TGACAG	AAT	TC	CAAT	TGCTTGC	7375
Db	5700	ACCTCCACGAGAA	AAGAG	CGGTTG	CTCTGT	CAG	AATCT	AC	CGTG	TCTTCTGCTTGGC	5759
Qy	7376	GGAGCTCGCCACT	TAAGAC	CTT	CGGTAG	CT	CGG	CGCTG	TAT	AGCGGAC	7435
Db	5760	GGAGCTCGCCACA	AAAGAC	CTT	CGG	AGCT	CCG	AA	TCG	TG	5819
Qy	7436	GACCGCCCTTCT	GACCTG	CGCT	CCGAC	GCGGTG	ACAA	AGGAT	TC	CGACGTTG	7495
Db	5820	AACGGCTCTCTG	AC	CAGCCCT	CCG	ACG	AGCGG	AC	CGCGG	ATCCG	5879
Qy	7496	CTCTCTCATG	CCCCCTT	GAA	GGGAG	CGCGGG	AC	CCCG	ATCT	CAGCG	7555
Db	5880	CTCTCTCATG	CCCCCTT	GAG	GGGAG	CGCGGG	AT	CCCG	ATCT	CAGCG	5939
Qy	7556	GTCTACCGTGAG	TGAGAG	GCTAG	TAGG	ATG	TCGT	CTG	CTCA	TGCTCCTAT	7615
Db	5940	GTCTACCGT	TAGCGAG	GAGGCT	TAG	TAGG	ACG	TCG	TGCT	CGGATG	5999
Qy	7616	GACAGGCGCCTG	AT	CAGCC	AT	CGCTG	CGGAG	AAAG	T	AAAGCTG	7675
Db	6000	GACAGGCGCCTG	AT	CAGCC	AT	CGCTG	CGGAG	AAAG	T	AAAGCTG	6059
Qy	7676	GAGCAACTCTTT	TG	TGGT	CACCA	CA	ATG	GT	TAC	CGCCAA	7735
Db	6060	GAGCAACTCTTT	TG	TGGT	CACCA	CA	ATG	GT	TAC	CGCCAA	6119
Qy	7736	CCTCGG	CAGAA	AGGT	CTAC	TTT	TGA	CAG	AT	TC	7795
Db	6120	CCTCGG	CAGAA	AGGT	CTAC	TTT	TGA	CAG	AT	TC	6179
Qy	7796	CGTACT	CAAG	GAGAT	G	AGGCG	AGGCG	TCC	ACAG	TTT	7855
Db	6180	CGTACT	CAAG	GAGAT	G	AGGCG	AGGCG	TCC	ACAG	TTT	6239
Qy	7856	GGAGG	CGCTG	CAAG	CTG	AC	CGCC	CC	CC	CA	7915
Db	6240	GGAGG	CGCTG	TAAG	CTG	AC	CGCC	CC	CC	CA	6299
Qy	7916	GGAGCT	CCG	GAAC	CT	TAT	CC	AG	CAG	GGCG	7975
Db	6300	GGAGCT	CCG	GAAC	CT	TAT	CC	AG	CAG	GGCG	6359

QY	7976	GCTGGAAGACACTGAAACACCAATTGACACCACCATCATGCGCAAAAAGTGAGGTTTTCTG	8035
Db	6360	GCTGGAAGACACTGAGACACCAATTGACACCACCATCATGCGCAAAAAGTGAGGTTTTCTG	6419
QY	8036	CGTCCAAACAGAGAAGGGAGCGCGAAGCCAGCTCGCCTTATCGTATTTCCAGACCTGGG	8095
Db	6420	CGTCCAAACAGAGAAGGGGGCGCGAAGCCAGCTCGCCTTATCGTATTTCCAGATTTGGG	6479
QY	8096	AGTTCTGTGATGCGAGAAGATGGCCCTTTACGACGTGGTCTCCACCCTTCCTCAGGCCGT	8155
Db	6480	GGTTCTGTGTGCGAGAAATAGCCCTTTACGATGTGGTCTCCACCCTCGCTCAGGCCGT	6539
QY	8156	GATGGGCTCCTCATACGGATTTTCAATACTCCCCAAGCAGCGGGTCCAGTTCTCTGGTAA	8215
Db	6540	GATGGGCTCTTCATACGGATTCCAACTACTCTCTGGACAGCGGGTCCAGTTCTCTGGTAA	6599
QY	8216	TACCTGGAATCAAAGAAATGCCCTTATGGGCTTCTCATATGACACCGCGCTGTTTGACTC	8275
Db	6600	TGCCTGAAAGCGAAGAAATGCCCTTATGGGCTTCGCATATGACACCGCGCTGTTTGACTC	6659
QY	8276	AACGGTCACTGAGAGTGACATTCGTCTGAGGAGTCAATTTACCAATGTTGTGACTTGGC	8335
Db	6660	AACGGTCACTGAGAATGACATTCGTGTGAGGAGTCAATCTACCAATGTTGTGACTTGGC	6719
QY	8336	CCCGGAGGCAGACAGGCCATAAGTTCGCTCAGAGAGCGGCTTTTACATCGGGGGTCCCC	8395
Db	6720	CCCGAAGCCAGACAGGCCATAGGTCGCTCAGAGAGCGGCTTTTACATCGGGGGGCCCC	6779
QY	8396	GACTAACTCAAAAGGCGAGAACTCGCGTTATTCGCCGTCGCGCAAGTGCGGTGCTGAC	8455
Db	6780	GACTAACTTCAAAAGGCGAGAACTCGCGCTATTCGCCGTCGCGCGAGCGGTGACTGAC	6839
QY	8456	GACTAGTGGGTAATACCTTCACATGTTACTTTGAAGGCCACTGACGCGCTGTCGAGCTGC	8515
Db	6840	GACACAGTGGGTAATACCTTCACATGTTACTTTGAAGGCCGCTGCGGCGCTGTCGAGCTGC	6899
QY	8516	AAAGCTCCAGGACTGCACGATGTCGTGAAACGGAGACGACCTTGTCTGTATCTGTGAAAG	8575
Db	6900	GAAGCTCCAGGACTGCACGATGTCGTATCGGAGACGACCTTGTCTGTATCTGTGAAAG	6959
QY	8576	CGCGGGAACCCAGGAGGATCGCGCGGCCCTTACGAGCGCTTACGCGAGGCTATGACTAGTA	8635
Db	6960	CGCGGGACCCAGAGGACGAGCGAGGCTTACGCGGCCCTTACGCGAGGCTATGACTAGATA	7019
QY	8636	TTCCGCGCCCCCGGGGATCCGCGCCCAACAGAAATACGACTCGAGCTGATAATCATATG	8695
Db	7020	CTCTGCGCCCCCTGGGGACCGGCCAAACAGAAATACGACTTGGAGTTGATAACATCATG	7079
QY	8696	TTCTCTCAATGTGTCACTCGCGACGATGATCTGGCAAAAGGTTACTACTCTCACCCG	8755
Db	7080	CTCTCTCAATGTGTCACTCGCGACGATGATCTGGCAAAAGGTTACTACTCTCACCCG	7139
QY	8756	TGACCCCAACACCCCTTGCACGGGCTCGGTGGGAGACAGCTAGACACACTCCAATCAA	8815
Db	7140	TGACCCCAACACCCCTTGGCGGGCTGGGTGGGAGACAGCTAGACACACTCCAATCAA	7199
QY	8816	CTCTTGCTAGGCAATATCATATGATGGCCACCCCTATGGSCAAAGGATGATTTCTGAT	8875
Db	7200	TTCTTGCTAGGCAACATCATATGATGGCCACCCCTTGTGGSCAAAGGATGATTTCTGAT	7259
QY	8876	GACTCACTTTTCTCCATCTCTTAGCTCAAGAGCAACTTTGAAAGGCCCTGAGTTGTCA	8935
Db	7260	GACTCACTTTTCTCCATCTCTTAGCTCAGGAAACAACTTTGAAAGGCCCTTAGTTGTCA	7319
QY	8936	GATCTACGGGCTTGCTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTGAAACGACT	8995
Db	7320	GATCTACGGGCTTGCTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTGAAACGACT	7379
QY	8996	CCATGGTCTTAGCGCATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGC	9055
Db	7380	CCATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGC	7439

Qy	9056	TTCATGCTCCAGAAACTTGGGGTACCA	CCCTTGGCAACTGGAGACATCGGCCAGAAG	9111
Db	7440	TTCATGCTCCAGAAACTTGGGGTACCG	CCCTTGGCAGATCTGGGCCAGAAG	7499
Qy	9116	TGTCGCGCTAAAGTACTGTCCAGGGG	GGGAGGCGCCACTTTGTGGCAGATACCTCTT	9175
Db	7500	TGTCGCGCTAGGCTACTGTCCAGGGG	GGGAGGCTGCCACTTTGTGGCAGATACCTCTT	7559
Qy	9176	TAACTGGGCGAGTAAGGACCAAAGCTT	AAACTCACTCCAAATCCCGGCTCCCAAGTTGGA	9235
Db	7560	CAACTGGGCGAGTAAGGACCAAAGCTT	AAACTCACTCCAAATCCCGGCTCCCAAGTTGGA	7619
Qy	9236	CTTGCTGCGCTGGTTCGTGCTGGTTAC	AGCGGGGAGACATATACAGCCTGTCTCG	9295
Db	7620	TTTATCCAGCTGGTTCGTGCTGGTTAC	AGCGGGGAGACATATACAGCCTGTCTCG	7679
Qy	9296	TGCCCAACCCGCTGGTTCGTGCTGGTT	CCGCTTCTGTAGGGGTAGGCATTTA	9355
Db	7680	TGCCCAACCCGCTGGTTCATGTGGTGC	TACTCTCTACTTTCTGTAGGGGTAGGCATCTA	7739
Qy	9356	CGTCTCCCCAACCGGATGAACGGGGAG	CTAAACCACTCCAGGCGCTTAAGCC-----	9405
Db	7740	TCTACTCCCCAACCGGATGAACGGGGAG	CTAAACCACTCCAGGCGCTTAAGCC-----	7799
Qy	9406	ATTCTCGTTTTTTTTTTTTTTTTTTTT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	9465
Db	7800	TTTTTCCCTTTTTTTTTTTTTTTTTTT	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	7859
Qy	9466	TCTTTTTTCCTTTCTTTTTTCCCTTC	TTTAAATCGTGGCTCCATCTTAGCCCTAGTCA	9525
Db	7860	TTTTTTCCTTTCTTTTTTCCCTTTCT	TTTCCCTTGGTGGCTCCATCTTAGCCCTAGTCA	7919
Qy	9526	CTAGCTGTGAAAGGTCGGTGAGCCGAT	GACTGCAGAGAGTGTGTGATCTGGCCTCTCTG	9585
Db	7920	CTAGCTGTGAAAGGTCGGTGAGCCGCT	TGACTGCAGAGAGTGTGTGATCTGGCCTCTCTG	7979
Qy	9586	CAGATCATGT 9595		
Db	7980	CAGATCAAGT 7989		
RESULT 15				
US-11-111-686-3				
; Sequence 3, Application US/11111686				
; Publication No. US2005026021A1				
; GENERAL INFORMATION:				
; APPLICANT: ANADYS Pharmaceuticals, Inc.				
; APPLICANT: Bichko, Vadim				
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFIC				
; FILE REFERENCE: 0342/1H3950S1				
; CURRENT APPLICATION NUMBER: US/11/111,686				
; CURRENT FILING DATE: 2005-04-20				
; PRIOR APPLICATION NUMBER: US/10/005,469				
; PRIOR FILING DATE: 2002-04-18				
; PRIOR APPLICATION NUMBER: US 60/245,866				
; PRIOR FILING DATE: 2000-11-07				
; NUMBER OF SEQ ID NOS: 14				
; SOFTWARE: PatentIn version 3.1				
; SEQ ID NO 3				
; LENGTH: 7995				
; TYPE: DNA				
; ORGANISM: Artificial Sequence				
; FEATURE:				
; OTHER INFORMATION: HCV Replicon RNA from cell line HCVR8				
US-11-111-686-3				
Query Match 55.3%; Score 5302.2; DB 7; Length 7995;				
Best Local Similarity 91.3%; Pred. No. 0;				
Matches 5652; Conservative 0; Mismatches 528; Indels 13; Gaps 2;				
Qy	3416	CTTTGGCCCATCATACGGCCTTACTCCCA	ACAAACGGGGGGCGTACTTGGTGCATCATCAC	3475
Db	1800	CATGGCGCTATTACGGCCTACTCCCAAC	AGACGCGGCGCTACTTGGCTGCATCATCAC	1859

Query Match	55.3%	Score 5302.2	DB 7	Length 7995
Best Local Similarity	91.3%	Pred. No. 0		
Matches 5652	Conservative	0	Mismatches 528	Indels 13
Gaps				
QY	3416	CTTTGGCGCCATCATCGCGCTACTCCCAACAAACGCGGGGGGTACTTGGTTGCATCATCAC	34175	
DB	1800	CATGCGCCCTATTACGCGCTACTCTCCACAGACGCGAGGCGCTACTTGGCTGCATCATCAC	1859	

Qy 3476 TAGCTCAGCGCGGAGAAACAGGTCGAAGGGGAGGTTCAAGTGGTTCCTACCGC 3535
Db |||||
Qy 1860 TAGCTCAGCGCGGAGACGACAGGTCGAGGGGAGGTCCAAGTGGTCTCCACGC 1919
Db |||||
Qy 3536 AACAAATCTTCTCGCGACTGTGATCAACGGCGTGTGTGGAGCTGTACCAATGGCGC 3595
Db |||||
Qy 1920 AACAAATCTTCTCGCGACTGTGATCAATGGCGTGTGTGGAGCTGTCTATCATGTGTC 1979
Qy 3596 TGGCTCGAAGACCTTAGCGGTCCAAAGGTCCAATCACCCCAATGTACACCAATGTAGA 3655
Db |||||
Qy 1980 CGGCTCAAAGACCTTTCGCGGCCAAAGGGGCCAATCACCCAAATGTACACCAATGTGGA 2039
Qy 3656 CTGAGACCTCGTGGCTGGCAGGCGCCCCCGGGGCGGCTCCATGACACCAATGAGCTG 3715
Db |||||
Qy 2040 CCAGGACCTCGTGGCTGGCAGGCGCCCCCGGGGCGGCTCTTGTGACACCAATGACACTG 2099
Qy 3716 TGGAGCTCGGACCTTTACTTGGTCAAGAGACATGCTGATGTCAATCCGGTGGCGCGGCG 3775
Db |||||
Qy 2100 CGGAGCTCGGACCTTTACTTGGTCAAGAGGATGCGGATGTCAATCCGGTGGCGCGGCG 2159
Qy 3776 AGGCGACAGCGGGAAGTCTACTCTCCCGGAGCGGCTCTCTACCTGAAGGCTCCTC 3835
Db |||||
Qy 2160 GGGCGACAGCGGGGAGGCGCTACTCTCCCGGAGCGGCTCTCTACTTTGAAGGGCTCTTC 2219
Qy 3836 GGGTGGTCCATTTGCTTTCCTTTCGGGACAGTGTGGGCGTCTTTCGGGCTGTGTGTG 3895
Db |||||
Qy 2220 GGGCGGTCACTGTCTCTGCCCCGCGGACAGCTGTGGGCACTTTCGGGCTGTGGTGTG 2279
Qy 3896 CACCGGGGGTTCGGAAGGGTGGACTTTCATACCCGTTGAGTCTATGGAATACCAAT 3955
Db |||||
Qy 2280 CACCGAGGGTTCGGAAGGGTGGACTTTGTACCCGCTCGAGTCTATGGGAACCACTAT 2339
Qy 3956 GGGTCTCGGTCTTCAGAGCAACTCAACCCCCCGGCTGTACCGGAGACATTCCAAGT 4015
Db |||||
Qy 2340 GGGTCCCGGTCTTCAGGCAACTCTGTCCCTCCCGGCGTACCGGAGACATTCAGAGT 2399
Qy 4016 GGCACATCTGACGCTCTTACTGGCAGCGGCAAGAGCACCAAGTCCGGCTGGGTATGC 4075
Db |||||
Qy 2400 GGGCATCTACACGCCCCCTACTGTGAGCGCAAGAGCACTAAGGTGCCGCTGTGATGC 2459
Qy 4076 AGCCCAAGGGTACAAAGTGTCTGCTGAAACCGGCTCGGTGGCGGCACTTATAGGTTTGG 4135
Db |||||
Qy 2460 AGCCCAAGGGTATAAGGTGTGCTCTGAAACCGGCTCGGCGGCAACCTAGTTCGG 2519
Qy 4136 GGGTATATGTCGAAGGACACGGTATCGACCTTAACATCAGAACTGGGGTAAGAACCAT 4195
Db |||||
Qy 2520 GGGTATATGTCGAAGGACATGGTATCGACCTTAACATCAGAACTGGGGTAAGAACCAT 2579
Qy 4196 TACCACGGGCTCCATTTACGTAATCTCCACCTATGCGCAAGTTCCTTGGCGGCTGGCTG 4255
Db |||||
Qy 2580 CACACGGGTGCCCATCAAGTATCCACCTATGCGCAAGTTCCTTGGCGAGTTCCTTGGCGAGTGTG 2639
Qy 4256 TTCTGGGCGCTATGACATCAATATGTGATGAGTGCCACTCACTGACTGACTGACTAC 4315
Db |||||
Qy 2640 CTCTGGGCGCTATGACATCAATATGTGATGAGTGCCACTCACTGACTGACTGACTGACT 2699
Qy 4316 CATTTGGGCTCGGACAGTCTCGAACAAGGAGAGCGGTGTGAGGCGGCTGTGCTGT 4375
Db |||||
Qy 2700 TATCTTGGGCTCGGACAGTCTCGAACAAGGAGAGCGGTGTGAGGCGGCTGTGAGGCGGCTGTGCTGT 2759
Qy 4376 GCTCGGCAACGCTACACCTTCGGGATCGGTATCGGTGCGACACCCCAATATCGAGGAAT 4435
Db |||||
Qy 2760 GCTCGGCAACGCTACGCTTCGGGATCGGTACCGTGCACATCCCAATATCGAGGAGGT 2819
Qy 4436 AGGCTGTCCAAATGGAGAGATCCCCCTTCTATGGCAAGGCAATCCCCATTTAGGGCCAT 4495
Db |||||
Qy 2820 GGCTGTGTCAGCACTGGAGAAATCCCCCTTTATGGCAAGGCAATCCCCATTCGAGACCAT 2879
Qy 4496 CAAAGGGGGAGGCACTCATTTTCTGTCATTTTCCAGAGAAATGTGACGAGCTCGCGCG 4555
Db |||||
Qy 2880 CAAAGGGGGAGGCACTCATTTTCTGTCATTTCCAGAGAAATGTGATGAGCTCGCGCG 2939

Qy 4556 AAAGCTGACAGGCTCGGACTGAACGCTGTAGCATATTACCGGGCCTTGATGTGCTCGT 4615
Db |||||
Qy 2940 GAAGTGTCTCGGCTCGGACTCAATGCTGTAGCATATTACCGGGCCTTGATGTATCGT 2999
Db |||||
Qy 4616 CATACCGCTTATCGGAGACGTCGTTGCTGTGGCAACAGAGCTCTAATGACGGGTTTAC 4675
Db |||||
Qy 3000 CATACCAACTAGCGGAGACGTCATTTGCTGTAGCAACGAGCGCTCTAATGACGGGCTTAC 3059
Qy 4676 CGGCGATTTTGAATCTGAGTGTGCTGCAATACATGTGTACCCAGACAGTCGACTTCAG 4735
Db |||||
Qy 3060 CGGCGATTTGCACTGAGTGTGCAATACATGTGTACCCAGACAGTCGACTTCAG 3119
Qy 4736 CTGAGTCCCACTTCACTTGAAGCAGACCGTGTCCCAAGACGCGGTGTGCGGCTC 4795
Db |||||
Qy 3120 CTGAGACCGGACTTCACTTGAAGCAGACCGTGTCCCAAGACGCGGTGTGCGGCTC 3179
Qy 4796 GCAACGCGGAGGTAGAACTGGCAGGAGTGGAGTGGCATCTACAGGTTTGTGACTTCAG 4855
Db |||||
Qy 3180 GCAACGCGGAGGAGGACTGGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTTCAG 3239
Qy 4856 AGAACGCGCTCGGGCATGTTTCGATTTCTCGTCTCTGTGTGAGTGTATGACGCGGGCTG 4915
Db |||||
Qy 3240 AGAACGCGCTCGGGCATGTTTCGATTTCTCGTCTCTGTGTGAGTGTATGACGCGGGCTG 3299
Qy 4916 TGCTTGTATGAGCTCAACGCGCTGAGACCTCGGTTAGGTTGCGGCTTACCTAAATAC 4975
Db |||||
Qy 3300 TGCTTGTACGAGCTCAACGCGCGGAGACCTCAGTTAGGTTGCGGCTTACCTAAACAC 3359
Qy 4976 ACAGGTTGCGGCTGTGCCAGGACCATCTGGAGTTCCTGGAGAGCGTCTTTCAGAGGCT 5035
Db |||||
Qy 3360 ACCAGGTTGCGGCTGTGCCAGGACCATCTGGAGTTCCTGGAGAGCGTCTTTCAGAGGCT 3419
Qy 5036 CACCCACATAGTGCACCTTCTGTCTCCAGACTAAACAGCAGGAGACAACTTTCCTTA 5095
Db |||||
Qy 3420 CACCCACATAGACGCGGCTTCTTGTCTCCAGACTAAGCAGGAGACAACTTCCCTTA 3479
Qy 5096 CCTGTTGGCATATCAAGCTACAGTGTGCGCAGGCTCAAGCTCCACTCCATCGTGGGA 5155
Db |||||
Qy 3480 CCTGTTAGCATACAGGCTACGCTGTGGCAGGCTCAGGCTCCACTCCATCGTGGGA 3539
Qy 5156 CCAATGTGGAAGTGTCTCATACGGCTGAAACCTTACCTGACGCGGCGCAACCGCTGTCT 5215
Db |||||
Qy 3540 CCAATGTGGAAGTGTCTCATACGGCTTAAAGCTTACCTGACGCGGCGCAACCGCTGTCT 3599
Qy 5216 GTATAGCTAGGAGCGCTCCAAATAGGTCATCTCACACACCCCAATACTAATAATACAT 5275
Db |||||
Qy 3600 GTATAGCTAGGAGCGCTTCAAAACGAGGTTACTTACACACACCCCAATACTAATAATACAT 3659
Qy 5276 CATGGCATGATGTGCGCTGACCTGAGGTCGTCACTAGCACCTGGGTGTCTGAGGCGG 5335
Db |||||
Qy 3660 CATGGCATGATGTGCGCTGACCTGAGGTCGTCACTAGCACCTGGGTGTCTGAGGCGG 3719
Qy 5336 AGTCTTGGAGCTTTGGCGGCTACTGCTGACGAGCGAGTGTGTGTCATTTGGGCGAG 5395
Db |||||
Qy 3720 AGTCTTGGAGCTTTGGCGGCTACTGCTGACAAACGAGCGTGTGTGTCATTTGGGCGAG 3779
Qy 5396 GATCATCTTGTTCGGGAGCGGCTGTCTCCGAGGGAAGTCTCTTACCGAGGCT 5455
Db |||||
Qy 3780 GATCATCTTGTTCGGGAGCGGCTCATCTTCCGAGAGGAAAGTCTTTCACCGGAGGT 3839
Qy 5456 CGATGAGTGAAGAGTGTGCTTCACTTCTTATCATCGAGAGGGAATGCACTGCTCGC 5515
Db |||||
Qy 3840 CGATGAGTGAAGAGTGTGCTTCACTTCTTATCATCGAAGGGAATGCACTGCTCGC 3899
Qy 5516 CGAGCAATTCAGCAAAAGGCGCTCGGTTGTGCAACGCGGCAACCAAGAGCGGAGGCG 5575
Db |||||
Qy 3900 CGAACAATTCAAACAGAGGCAATCGGCTGTCTGCAACAGCCACCAAGCAAGCGGAGCG 3959
Qy 5576 TGCTGCTCCGCTGTGGAGTCCAAAGTGGCAGGCTTTCAGACCTTCTGGGCGAGACAT 5635
Db |||||
Qy 3960 TGCTGCTCCGCTGTGGAAATCCAAAGTGGCGGACCTTCCGAGGCTTCTGGGCGAGACAT 4019
Qy 5636 GTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATTCACCTCTGCTGGAACCC 5695

Db	4020	GTGGAAATTCATCAGCGGATCAATATTTAGCAGCTTGTCCACTCTCCCTGGCAACCC	4079	
Qy	5696	CGCGATAGCATCAATGATGGCAATTTACAGCTTCTATCACTAGCCCGCTCACCACCCAAAA	5755	
Db	4080	CGGGATAGCATCACTGATGGCAATTCACAGCTTCTATCAACAGCCCGCTCACCACCCAA	4139	
Qy	5756	CACCTCTCTGTTTAAATCTTTGGGGGATGGGTGGTGCCTCCAACTCGCTCTCCCGAGCG	5815	
Db	4140	TACCTCTCTGTTTAAATCTTTGGGGGATGGGTGGTGCCTCCAACTCTCTCCCGAGCG	4199	
Qy	5816	TGGGTGAGCTTCTGGGCGCGGATCGCGAGCGGCTGTGGCAGCATAGGCCCTG	5875	
Db	4200	TGCTTCTGCTTCTGATAGCGCGCATCGTGGAGCGGCTGTGGCAGCATAGGCCCTG	4259	
Qy	5876	GAAAGTGCTGCTGGACATCTTGGCGGCTATGGGCGAGGGTAGCGCGGCACTCGTGGC	5935	
Db	4260	GAAAGTGCTTGTGATATTTTGACAGGTTATGAGCAGGGGTGGCAGCGGCTCGTGGC	4319	
Qy	5936	CTTTAAAGGTATGAGCGGAGGTGCCCTCCACCGAGGACCTGGTCAACTTACTCCCTGC	5995	
Db	4320	CTTTAAAGGTATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGC	4379	
Qy	5996	CATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6055	
Db	4380	TATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4439	
Qy	6056	CGTGGCGCGGAGAGGGGCTGTGAGTGAGTGAACCGGCTCATAGCGTTCCGCTCGCG	6115	
Db	4440	CGTGGCGCGGAGAGGGGCTGTGAGTGAGTGAACCGGCTCATAGCGTTCCGCTCGCG	4499	
Qy	6116	GGGTAAACACGCTCTCCCTACGCACTATGTGCTGAGAGCGAGCTGCGAGCAGTGTAC	6175	
Db	4500	GGGTAAACACGCTCTCCCTACGCACTATGTGCTGAGAGCGAGCTGCGAGCAGTGTAC	4559	
Qy	6176	TCAGATCTCTAGCTTACATCACTCACTGCTGAACGGCTCCACAGTGGATTAA	6235	
Db	4560	TCAGATCTCTAGCTTACATCACTCACTGCTGAACGGCTCCACAGTGGATTAA	4619	
Qy	6236	TGAGGAGCTCTACGCGCATCTCGGCTCGTGGCTAAGGATGTTTGGGATGGATATG	6295	
Db	4620	CGAGGAGCTCTACGCGCATCTCGGCTCGTGGCTAAGGATGTTTGGGATGGATATG	4679	
Qy	6296	CACGGTGTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCTGCGCGGGTTACCGGG	6355	
Db	4680	CACGGTGTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCTGCGCGGGTTACCGGG	4739	
Qy	6356	AGTCCCTTTCTGTGTCATGCGGATCAAGGGAGTCTGCGGGGGGACGGCATCAT	6415	
Db	4740	AGTCCCTTTCTGTGTCATGCGGATCAAGGGAGTCTGCGGGGGGACGGCATCAT	4799	
Qy	6416	GCAAAACACCTGCGGAGCAGATCGCGGACATGT--CAAAAACGGTTCCAT	6472	
Db	4800	GCAAAACACCTGCGGAGCAGATCGCGGACATGTGAAAACAAACGGTTCCAT	4859	
Qy	6473	GAGGATCGTAGGGCTTAGAACCTTCAGCAACACGCTGGCAGGAAGCTTCCCATCAACGC	6532	
Db	4860	GAGGATCGTAGGGCTTAGAACCTTCAGCAACACGCTGGCAGGAAGCTTCCCATTAACGC	4919	
Qy	6533	ATACACCAACGGGACCTTGCACACCTCCCGCGGCGCAACTATTTCCAGGGCGCTATGGCG	6592	
Db	4920	GTAACACAGGGCGGCTGACCGCTCCCGCGGCGCAAAATTTATTTCTAGGGCGCTGTGGCG	4979	
Qy	6593	GGTGGCTGCTAGAGGATAGTGGAGGTTACCGGTGTGGGGATTTCCACTACGTGACGGG	6652	
Db	4980	GGTGGCTGCTAGAGGATAGTGGAGGTTACCGGGTGGGGATTTTCCACTACGTGACGGG	5039	
Qy	6653	CATGACCACTCACAACGTAAGTGCCCATGCCAGGTTCCGGCGGCGGAAATTTCTTACGGA	6712	
Db	5040	CATGACCACTCACAACGTAAGTGCCCATGCCAGGTTCCGGCGGCGGAAATTTCTTACAGA	5099	
Qy	6713	GGTGGATGGAGTGGGTTGACAGGTACGCTCCGGCGGTGCAAACTCTTCTTACGGGAGGA	6772	

Db	5100	AGTGGATGGGTGCGGTTGTCACAGGTACGCTCCAGGTGTCAAAACCCCTCTTACGGAGGA	5159	
Qy	6773	CGTCAAGTTCCAGGTGCGGCTCAACAATATCTTGGTGGGTGCGAGCTCCCATGGAGCC	5832	
Db	5160	GGTCAATTTCTGGTGGGCTCAATCAATACCTGTTGGGTGCGAGCTCCCATGGAGCC	5219	
Qy	6833	CGAACCGGAGCTAAACAGGTGCTTACCTTCCATGCTACCGGATCCCTCCACATATACAGAGA	6892	
Db	5220	CGAACCGGAGCTAAACAGGTGCTTACCTTCCATGCTACCGGATCCCTCCACATATACAGAGA	5279	
Qy	6893	GACGGCTAAGCGTAGCTGGCTAGAGGCTTCCCGCTCTTTAGCAGCTCATFACGTAG	6952	
Db	5280	GACGGCTAAGCGTAGCTGGCTAGAGGCTTCCCGCTCTTTAGCAGCTCATFACGTAG	5339	
Qy	6953	CGAGTGTCTGGCTTCTTTGAGGGGACATGCTACCTACCCAGCTCCCGGAGCC	7012	
Db	5340	CGAGTGTCTGGCTTCTTTGAGGGGACATGCTACCTACCCAGCTCCCGGAGCC	5399	
Qy	7013	TGACCTCATCGAGGCGCAACCTCTTGTGGCGGAGGAGATGGGCGGAAACATCACTCGCGT	7072	
Db	5400	TGACCTCATCGAGGCGCAACCTCTTGTGGCGGAGGAGATGGGCGGAAACATCACTCGCGT	5459	
Qy	7073	GGAGTCAGAGAAATAGGTAGTAAATTTCTGACCTTTTCGAACCGCTTCAACGGAGGGGA	7132	
Db	5460	GGAGTCAGAGAAATAGGTAGTAAATTTCTGACCTTTTCGAGCGCTCCAAGCGGAGGA	5519	
Qy	7133	TGAGAGGAGATATCCGTCGCGGAGATCTCGGAAATCCAGGAAGTTCCCTCAGC	7192	
Db	5520	TGAGAGGAGATATCCGTCGCGGAGATCTCGGAAATCCAGGAAGTTCCCTCAGC	5579	
Qy	7193	GTTGCCATATGGGACGCGCGGACTACAACTCTCCACCTGTAGAGTCTTGGAGGAGCC	7252	
Db	5580	GATGCCATATGGGACGCGCGGATTTACAACTCTCCACCTGTAGAGTCTTGGAGGAGCC	5639	
Qy	7253	GGACTACGTCCTCCGCTGGTACACGGATGCCATTTGCCACCTTACCAAGGCTCTTCCAAT	7312	
Db	5640	GGACTACGTCCTCCGCTGGTACACGGATGCCATTTGCCACCTTACCAAGGCTCTTCCAAT	5699	
Qy	7313	ACCACCTCCAGGAGAAAGGAGCGGTTGTCTGACAGAAATCCAAATGTGTCTTCTGCTT	7372	
Db	5700	ACCACCTCCAGGAGAAAGGAGCGGTTGTCTGACAGAAATCCAAATGTGTCTTCTGCTT	5759	
Qy	7373	GGCGGAGCTCGGCACTTAAGACCTTCGGTACCTCGGATCGTGGCGCTTGTATAGCGCAC	7432	
Db	5760	GGCGGAGCTCGGCACTTAAGACCTTCGGTACCTCGGATCGTGGCGCTTGTATAGCGCAC	5819	
Qy	7433	GGCGGAGCTCGGCACTTCGACCTCCGACGAGCGGTGACAAAGGATCCGACGTTGAGTC	7492	
Db	5820	GGCAACGGCTCTCTTGACAGGCGCTCCGACGAGCGGATCCGCGGATCCGACGTTGAGTC	5879	
Qy	7493	GTACTCTCCATCCCGCTTGAAGGGAGCGGGGACCCCGATCTCAGCGAGCGGTC	7552	
Db	5880	GTACTCTCCATCCCGCTTGAAGGGAGCGGGGATCCCGATCTCAGCGAGCGGTC	5939	
Qy	7553	TTGGTCTACCGTAGGAGGAGCTAGTAGGATGTCTGCTGCTGCTCAATGTCTTATAC	7612	
Db	5940	TTGGTCTACCGTAGGAGGAGCTAGTAGGAGCTGCTGCTGCTGCTCAATGTCTTATAC	5999	
Qy	7613	GTGAGACGGCGCTGATACGCTGCTGGGAGGAAAGTAAGTGGCCCATCAACCC	7672	
Db	6000	ATGAGAGGGCGCTGATACGCTGCTGGGAGGAAAGTAAGTGGCCCATCAATGC	6059	
Qy	7673	GTTGAGCAACTCTTGTGCTGCTCAACAAATGGTCTACGCGCAACATCCCGGAGCG	7732	
Db	6060	ACTGAGCAACTCTTGTGCTGCTCAACAAATGGTCTACGCGCAACATCCCGGAGCG	6119	
Qy	7733	AAGCTCTCGGAGAGGAGTCACTTTTGAAGATTCAGATTCAGATTCATTAACCG	7792	
Db	6120	AAGCTCTCGGAGAGGAGTCACTTTTGAAGATTCAGATTCAGATTCATTAACCG	6179	
Qy	7793	GGAGTACTCAGGAGATGAAGCGGAGGCTCCACAGTTAAGCTAAGCTTCTTATCTAT	7852	
Db	6180	GGAGTACTCAGGAGATGAAGCGGAGGCTCCACAGTTAAGCTAAGCTTCTTATCTAT	6239	

Qy	7853	AGAGGAGCTGCAAGCTGACGCCCCCACAATTCGGCCAAATCCAAATTTGGCTATGGGC	7912
Db	6240	CGAGGAGCCTGTAACTGACGCCCCCACAATTCGGCCAGATCTAAATTTGGCTATGGGC	6299
Qy	7913	AAAGSAGCTCCGGAACTTATCCAGACGGCCGTTAAACACATCCGCTCCGTTGGGAGGA	7972
Db	6300	AAAGSAGCTCCGGAACTTATCCAGACGGCCGTTAAACACATCCGCTCCGTTGGGAGGA	6359
Qy	7973	CTTGCTGAAGACACTGAACACCAATTTGACACCAACCATCATGGCAAAAAGTGAGTTTT	8032
Db	6360	CTTGCTGAAGACACTGAAGACCAATTTGACACCAACCATCATGGCAAAAATGAGTTTT	6419
Qy	8033	CTGCGTCCAAACAGAGAGGGAGCGCGCAAGCCAGCTCGCCTTATCGTATTTCCAGACCT	8092
Db	6420	CTGCGTCCAAACAGAGAGGGAGCGCGCAAGCCAGCTCGCCTTATCGTATTTCCAGATTT	6479
Qy	8093	GGGAGTTGCTGTATGCGAGAGATGGCCCTTTTACGACGTGTCTCCACCTTCTCTCAGGC	8152
Db	6480	GGGAGTTGCTGTGCGAGAGAAATGGCCCTTTTACGATGTGTCTCCACCTTCTCTCAGGC	6539
Qy	8153	CGTATGGGCTCCTCATAGGATTTCAATATCTCCCCCAAGCAAGCGGTGCGAGTTCCCTGGT	8212
Db	6540	CGTATGGGCTCCTCATAGGATTTCAATATCTCTCGACAGCGGGTTCGAGTTCCCTGGT	6599
Qy	8213	GAATACCTGGAAATCAAGAAATGCCCTATGGGCTTCTCATATGACACCGCTGTTTTGA	8272
Db	6600	GAATGCCCTGGAAAGCGAAGAAATGCCCTATGGGCTTCTCGATATGACACCGCTGTTTTGA	6659
Qy	8273	CTCAACGGTCACTGAGAGTGACATTCGTGTGGAGGTCAATTTACCAATGTTGTGACTT	8332
Db	6660	CTCAACGGTCACTGAGATGACATTCGTGTGGAGGTCAATCTACCAATGTTGTGACTT	6719
Qy	8333	GGCCCCGAGGCGACAGGCGCATTAAGGTTCGTCTACAGAGGGCTTTACATCGGGGGTCC	8392
Db	6720	GGCCCCGAGGCGACAGGCGCATTAAGGTTCGTCTACAGAGGGCTTTACATCGGGGGGCC	6779
Qy	8393	CCTGACTAATCTAAAGGGCAGAACTCGGGTTATCGCGGTTCGCGCAAGTGGCGTCT	8452
Db	6780	CCTGACTAATCTAAAGGGCAGAACTCGCGGTTCGCGGTTCGCGCAAGTGGCGTCT	6839
Qy	8453	GACGACTAGCTCGGTATATACCTTCACATGTTTACTTTGAAGGCCACTGACGCGCTGCGAGC	8512
Db	6840	GACGACCACTCGGTATATACCTTCACATGTTTACTTTGAAGGCCCTGCGCGCTGCGAGC	6899
Qy	8513	TGCAAGCTCCAGGACTGACGATGCTGTGTAAACGAGACGACCTTGTGTTATCTGTGA	8572
Db	6900	TGCGAAGCTCCAGGACTGACGATGCTGTGTAAACGAGACGACCTTGTGTTATCTGTGA	6959
Qy	8573	AAGCGGGGNAACCCAGGAGATGCGCGGCCCTTACAGGCTTTCAGCGAGGCTATGACTAG	8632
Db	6960	AAGCGGGGNAACCCAGGAGATGCGCGGCCCTTACAGGCTTTCAGCGAGGCTATGACTAG	7019
Qy	8633	GTATTCGCGCCCGCCCGGGATCCGCCCCCAACAGAAATACGACCTGGAGCTGATAACATC	8692
Db	7020	ATACTCTGCCCCCTGGGGAACCGCCCCCAACAGAAATACGACTTGGAGTTGATAACATC	7079
Qy	8693	ATGTTCTCTCAATGTGTAGTCCGCGACGATGCTGTGCAAAAGGGTATACTACCTCAC	8752
Db	7080	ATGTTCTCTCAATGTGTAGTCCGCGACGATGCTGTGCAAAAGGGTATACTACCTCAC	7139
Qy	8753	CGGTGACCCCAACCCCTTTCGACGGGCTCGGTGGGAGACAGCTAGACACACTCCCAAT	8812
Db	7140	CGGTGACCCCAACCCCTTTCGACGGGCTCGGTGGGAGACAGCTAGACACACTCCCAAT	7199
Qy	8813	CAACTCTTGCTAGGCAATATCATGATGATGCGGCCCACTTATGGCAAGGATGATCT	8872
Db	7200	CAATTCCTGGCTAGGCAACATCATCATGATGCGGCCCACTTGTGGCAAGGATGATCT	7259
Qy	8873	GATGACTCATCTTTTCTCCATCCTTCTAGCTCAAGAGCAACTTTGAAAAAGCCCTGGAATG	8932
Db	7260	GATGACTCATCTTTCTCCATCCTTCTAGCTCAGGAACTTTGAAAAAGCCCTAGATTTG	7319

Qy	8933	TCAGATCTACGGGCTTGTCTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTGAACG	8992
Db	7320	TCAGATCTACGGGCTTGTCTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTGAACG	7379
Qy	8993	ACTCCATGGTCTTAGCGCATTTTACATCCACAGTTACTCTCCAGGTGAGATCAATAGGCT	9052
Db	7380	ACTCCATGGGCTTAGCGCATTTTACTCTCCATAGTTACTCTCCAGGTGAGATCAATAGGCT	7439
Qy	9053	GGCTTCATGCTCAGGAAACTTGGGGTACCACCTTCGGAACCTTGGAGACATCGGGCCAG	9112
Db	7440	GGCTTCATGCTCAGGAAACTTGGGGTACCACCTTCGGAAGCTTGGAGACATCGGGCCAG	7499
Qy	9113	AAAGTGTCCGGCTAAAGCTACTGTCCAGGGGGGAGCGCCCACTTTGTGGCAGATACCT	9172
Db	7500	AAAGTGTCCGGCTAAAGCTACTGTCCAGGGGGGAGCGCGCTGCCACTTTGTGGCAAGTACCT	7559
Qy	9173	CTTTAACTGGGCGAGTAAAGCAAGCTTAACTCACTCTCAATCCCGCGCGGTCCCAGCT	9232
Db	7560	CTTTCAACTGGGCGAGTAAAGCAAGCTTCAAACTCACTCAATCCCGCGGTGCCAGTT	7619
Qy	9233	GGACTTGTCTGGCTGGTTCGCTCGCTGTACAGCGGGGAGACATATATCACAGCCTGTC	9292
Db	7620	GGATTTATCAGCTGGTTCGCTGTGTACAGCGGGGAGACATATATCACAGCCTGTC	7679
Qy	9293	TCGTGCCCGACCCCGCTGGTTTCGTTGTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCAT	9352
Db	7680	TCGTGCCCGACCCCGCTGGTTTCATGTGTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCAT	7739
Qy	9353	TTACTCTCTCCCAACCGATGAACGGGGAGTAAACCACTCCAGGCCCTTAAGCC-----	9405
Db	7740	CTATCTACTCCCAACCGATGAACGGGGAGTAAACCACTCCAGGCCCAATAGGCCATCCTG	7799
Qy	9406	---ATTTCCTGTT	9462
Db	7800	TTTTTTTTCCCTTT	7859
Qy	9463	CTTTCTTTTTTTTCCCTTTTCCCTTTTAAATGGTGGCTCCATCTTTAGCCCTTAGTCA	9522
Db	7860	TTTTTTTTTCTCTTTTTTTTTTCTTTTCTTTCTTTGGTGGCTCCATCTTTAGCCCTTAGTCA	7919
Qy	9523	CGGCTAGCTGTGAAAGGTCCGCTGAGCCGCAATGACTGACAGAGTGTGTGATCTGGCCTCT	9582
Db	7920	CGGCTAGCTGTGAAAGGTCCGCTGAGCCGCTTGTACTGACAGAGTGTGTGATCTGGCCTCT	7979
Qy	9583	CTGCAGATCATGT 9595	
Db	7980	CTGCAGATCAAGT 7992	

Search completed: December 12, 2005, 06:20:35
Job time : 543 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 11, 2005, 04:29:02 ; Search time 22235 Seconds
(without alignments)
20189.883 Million cell updates/sec

Title: US-09-662-454-4
Perfect score: 9595
Sequence: 1 gccagccccctgatggggc.....ggcctctgcagatcatgt 9595

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_btc: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gss1: *
10: gb_gss2: *
11: gb_gss3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	419.2	4.4	980	10	CZ934124 251577 To
2	375	3.9	824	10	CZ990751 207976 To
C 3	354.8	3.7	849	10	CZ990300 207380 To
C 4	318	3.3	769	10	CZ990230 207284 To
5	315.2	3.3	852	10	CZ990303 207384 To
6	287.2	3.0	817	10	CZ990242 207300 To
7	257.4	2.7	922	10	CZ990744 207968 To
8	139.6	1.5	891	10	CZ990535 207704 To
C 9	129	1.3	487	1	AV755731 AV755731
C 10	102.8	1.1	597	7	CV350260 MR2-SN000
C 11	98	1.0	492	1	AV758366 AV758366
C 12	79.4	0.8	832	7	CN648295 ILLUMIGEN
C 13	77.8	0.8	914	3	BQ232672 AGENCOURT
C 14	76.6	0.8	807	2	BG108244 602280129
C 15	76	0.8	796	10	CNS011GK Alu100238 Drosophil
C 16	75.6	0.8	278	3	BJ703053 BJ703053
C 17	74.8	0.8	548	3	BM812285 fx16f08.Y
C 18	74.6	0.8	269	3	BJ703875 BJ703875
C 19	74.4	0.8	562	5	BQ747212 UI-N-PAO
C 20	74.4	0.8	979	2	BF035653 601458622
C 21	74.2	0.8	304	3	BJ688640 BJ688640
C 22	74.2	0.8	656	4	BC033047 Mus muscu

C 23	74.2	0.8	848	6	CA758134
C 24	73.8	0.8	328	7	CO088881
C 25	73.8	0.8	400	7	CK621399
C 26	73.8	0.8	777	7	CN647221 ILLUMIGEN
C 27	73.8	0.8	1003	3	BM415727
C 28	73.6	0.8	276	8	DN374563
C 29	73.6	0.8	326	8	DN412367
C 30	73.4	0.8	566	6	CF754016
C 31	73.4	0.8	731	2	BG572984
C 32	73.4	0.8	759	7	CK809913
C 33	73.2	0.8	930	8	CV922381
C 34	73	0.8	202	2	BF814420
C 35	73	0.8	473	3	BJ686888
C 36	73	0.8	558	1	AL643867
C 37	72.8	0.8	183	2	BG996024
C 38	72.8	0.8	349	3	BJ687812
C 39	72.8	0.8	498	8	CK393284
C 40	72.8	0.8	664	2	BF144889
C 41	72.6	0.8	721	1	AM023294
C 42	72.6	0.8	868	7	CN647590
C 43	72.6	0.8	927	10	CL121325
C 44	72.4	0.8	187	8	DN652172
C 45	72.4	0.8	346	3	BJ701862

ALIGNMENTS

RESULT 1
LOCUS CZ934124 980 bp DNA linear GSS 11-AUG-2005
DEFINITION 251577 tomato EcorI BAC Library Lycopersicon esculentum genomic clone SL_EcorI0030115 5, genomic survey sequence.
ACCESSION CZ934124
VERSION CZ934124.1 GI:72264385
KEYWORDS GSS.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 980)
AUTHORS Mueller,L.A., Buel,S.R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J., Van Eck,J., and Stack,S.
TITLE BAC end sequencing from three Solanum lycopersicon libraries
JOURNAL Unpublished (2005)
COMMENT Other GSSs: 251576
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@sgn.cornell.edu
Plate: 30 row: I column: 15
Seq primer: T7
Class: BAC ends
High quality sequence start: 23
High quality sequence stop: 485.
Location/Qualifiers
1. 980
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Heinz 1706"
/db_xref="taxon:4081"
/clone="SL_EcorI0030115"
/lab_host="E. coli"
/clone_lib="Tomato EcorI BAC Library"
/note="Vector: unk; Site_1: EcoRI"

FEATURES
source
1. 980
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Heinz 1706"
/db_xref="taxon:4081"
/clone="SL_EcorI0030115"
/lab_host="E. coli"
/clone_lib="Tomato EcorI BAC Library"
/note="Vector: unk; Site_1: EcoRI"

ORIGIN
Query Match 4.4%; Score 419.2; DB 10; Length 980;
Best Local Similarity 84.7%; Pred. No. 1.5e-90;

Matches 505; Conservative 0; Mismatches 88; Indels 3; Gaps 3;

QY 3321 GGGGACATCATCTTGGGTCTACCGCTCTCCGCGAAGGGGAAGAGATATTTTGGGA 3380
|||||
Db 23 GGGGACATCATCTCGGGCCCTACCGGTCTCCGCGAAGGGGAAGAGATATTTCTAGGG 82
|||||

QY 3381 CCGGCTGATAGTCTCGAAGGGCAAGGGTGGCGACTCCTTTGCCGCCATCACGGCTTACTCC 3440
|||||
Db 83 CCGGCGGAGCGCTTGAAGACCTTTGGTGGCGACTCCTTTGCCGCCATCACGGCTTACTCC 142
|||||

QY 3441 CAACAAACCGGGGGCTACTTGGTGCATCATCACTAGCCTCACAGCGGGGACAAGAAC 3500
|||||
Db 143 CAACAGACCGGGGGCTACTTGGTGTATCATCACTAGCCTCACAGCGGGGACAAGAAC 202
|||||

QY 3501 CAGGTCGAAGGGGAGTTCAGGTGTTCTACCGCAACACATCTTCTCGCGACCTGC 3560
|||||
Db 203 CAGGTCGAAGGGGAGTTCAGGTGTTCTCACTGGGACGACGTCTTCTCGCGACCTGT 262
|||||

QY 3561 ATCAACGGGCTGTCTGGAAGTGTCTACCATGGCTGGCTCGAAGACCTTAGCCGGTCCA 3620
|||||
Db 263 GTCATGGGCTGTGTGACTGTCTTCCATGGCGCGCTCAAGACCTTAGCCGGCCTCA 322
|||||

QY 3621 AAAGTCCAAATCACCCAAATGTACACCAATGTAGACCTCGTGGCTGGCAGCG 3680
|||||
Db 323 AAAGGCCCAATCACCCAAATGTACACCAATGTAGACCAAGTGTGGCTGGCAGCG 382
|||||

QY 3681 CCCCCGGGGGGCTCCATGATACACCATGACGTGTGGAGCTCGGACCTTACTTGGTC 3740
|||||
Db 383 CCCCCGGGGGGCTCCATGATACACCATGACGTGTGGAGCTCGGACCTTACTTGGTC 442
|||||

QY 3741 ACAGACATGCTGATGCTATTCGGGTGCGCGGAGGGGACAGACGGGGAAGTCTACTC 3800
|||||
Db 443 ACAGACATGCTGATGCTATTCGGGTGCTCCAGCGGGGACAGACGGGGAAGTCTACTA 502
|||||

QY 3801 TCCCCCA-GGCGCGTCTCTACTGAAAGGCTCTCGGGTGT-CCATTGCTTTGCCCTT 3858
|||||
Db 503 TCCCCCATGCCGCTCTCTACTTGAAGGACTCTTCCGGTGTGCCACTGCTTCCGCCCT 562
|||||

QY 3859 CGGGGACGCTGTGGGC-GTCTTCGGGCTGTGTGTGACCGGGGGGTCCGAA 3913
|||||
Db 563 CAAGGTATTATGTTTGCATTCTCAGGGCTGCTTGAAGCTCTGTGGTGGCTCTAA 618
|||||

RESULT 2
CZ990751
LOCUS
DEFINITION
207976 Tomato MboI BAC Library Lycopersicon esculentum genomic
clone SL_MboI0131013 5, genomic survey sequence.

ACCESSION
CZ990751
VERSION
CZ990751.1
KEYWORDS
GI:72342396
SOURCE
Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 824)
Muller, L.A., Bueis, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,
Van Eck, J., and Stack, S.
BAC end sequencing from three Solanum lycopersicon libraries
Unpublished (2005)
Other_GSSs: 207975
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@sgn.cornell.edu
Plate: 131 row: O column: 13
Seq primer: T7
Class: BAC ends
High quality sequence start: 16

FEATURES
source
High quality sequence stop: 553.
Location/Qualifiers
1..824
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Heinz 1706"
/db_xref="taxon:4081"
/clone="SL_MboI0131013"
/lab_host="E. coli"
/clone_lib="Tomato MboI BAC Library"
/note="Vector: pBelOBAC11; Site_1: MboI"

ORIGIN
Query Match 3.9%; Score 375; DB 10; Length 824;
Best Local Similarity 72.4%; Pred. No. 8.7e-80;
Matches 500; Conservative 0; Mismatches 190; Indels 1; Gaps 1;

QY 3301 GAGCAGACACCGCTCGGTGGGACATCATCTTTGGGTCTACCGTCTTCCGCCGGAAGG 3360
|||
Db 2 CGCGATCGCCTTGGCGTGGGACATCATCGACGGCTTCCCGTCTCTGCCGTAGGG 61
|||

QY 3361 GGAAGGAGATATTTTGGGACCGCTGATAGTCTCGAAGGCAAGGTGGGACTCTCTTG 3420
|||
Db 62 CGCGAGAGATATCTGCTCGGCGCGCGATGGAATGGCTTCTAAGGGGTGGAGTTGTTGG 121
|||

QY 3421 CGCCATCACGGCTTACTCCCAACAAAGCGGGGGCTACTTGGTTCATCATCACTAGCC 3480
|||
Db 122 CGCCATCACGGCTGATGCCAGCAGAGGGGCTCTTGGGTGCATATATACAGCC 181
|||

QY 3481 TCAGAGCGGGGACAAGAACACAGGTGCAAGGGGAGTTTCAAGTGGTCTTACCCGCAAC 3540
|||
Db 182 TGACCGGGGACAAGAACACAGGTGGAGGTGAGGTTGAGATTGTGTCACTGCCGCC 241
|||

QY 3541 AATCTTCTCGGACCTGATCAACCGCGTGTGTGACTGTCTACCATGGCGTGCCT 3600
|||
Db 242 AGACCTTCTCGCAACCTGTATTAAACGGGGTGTGTGACCGTCTACCAACGGGCGGAA 301
|||

QY 3601 CGAAGACCTTAGCGGTCCAAAGTCCAAATGACCCAAATGTACACCAATGTAGACCTGG 3660
|||
Db 302 CAAGACCATCGCTCACCAAGGTCTCTTAATCCAGATGTACCAATGTGAGACCAAG 361
|||

QY 3661 ACTCTGCTGGCGAGCGGCCCGCGCGCTCCATGACACCATGCACTGCTGGCA 3720
|||
Db 362 ACTCTGAGGTGGCGCGCTCCCAAGGTGCCGCTCATTAACACCTGCACTTCCGGCT 421
|||

QY 3721 GCTCGACCTTACTTGTGTACGAGACATGTGTATGTCAATTCGGGTGCGCGCGAGGG 3780
|||
Db 422 CCTCGACCTTACTCTGTGTACGAGGACGCGCGATGTCAATTCGGGTGCGCGCGGGTG 481
|||

QY 3781 ACAGCAGGGGAAGTCTACTCTCCCCAGCGCGCTCTCTACTGAAAGGCTCTCTCGGGTG 3840
|||
Db 482 ACAGCAGGGGAGCTCTCTCGCCCCCGGCTATCTCTTACTTGAAGGGCTCTCTCGGGG 541
|||

QY 3841 GTCCATCTTTCCTTTCGGGCGACGCTGTGGGCGTCTTCGGGCTGTGTGTGACCC 3900
|||
Db 542 GCCCACTCTGTGCCCCCGGGGACACGCTGTATGCATATTTAGGGCCGCGGTGCCCCC 601
|||

QY 3901 GGGGGGTGCGAAGCGGTGGA-CTTCAATACCGTTGAGTCTATGGAACCTACCATCGGG 3959
|||
Db 602 GTGGAGTGTCTAAGCGCGTGGATTTTTCGCCCGGTAGAAAGCCCTTAAACACCATGAAG 661
|||

QY 3960 TCTCGGCTTTCACAGACAACCTCAACCCCCC 3990
|||
Db 662 GTCCCGGGTTCACCTGACCACTTCTTCTCCC 692
|||

RESULT 3
CZ990300/c
LOCUS
DEFINITION
207380 Tomato MboI BAC Library Lycopersicon esculentum genomic
clone SL_MboI0131C03 5, genomic survey sequence.

ACCESSION
CZ990300
VERSION
CZ990300.1
GI:72341945

KEYWORDS	GSS.
SOURCE	Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM	Lycopersicon esculentum
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
AUTHORS	1 (bases 1 to 849) Mueller,L.A., Bueis,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J., Van Eck,J. and Stack,S.
TITLE	BAC end sequencing from three Solanum lycopersicon libraries
JOURNAL	Unpublished (2005)
COMMENT	Other_GSSs: 207379 Contact: Lukas Mueller Tanksley Lab, Dept. of Plant Breeding Cornell University 251 Emerson Hall, Ithaca, NY 14853, USA Tel: 607-255-6557 Fax: 607-255-6683 Email: sgn-feedback@sgn.cornell.edu Plate: 131 row: C column: 3 Seq primer: T7 Class: BAC ends High quality sequence start: 25 High quality sequence stop: 464.
FEATURES	source 1..849 /organism="Lycopersicon esculentum" /mol_type="genomic DNA" /cultivar="Heinz 1706" /db_xref="taxon:4081" /clone="SL_MboI0131C03" /lab_host="E. coli" /clone_lib="Tomato MboI BAC Library" /note="Vector: pBelOBAC11; Site_1: MboI"
ORIGIN	Query Match 3.7%; Score 354.8; DB 10; Length 849; Best Local Similarity 71.8%; Pred. No. 7.6e-75; Matches 464; Conservative 0; Mismatches 182; Indels 0; Gaps 0;
Qy	3510 GGGAGGTTCAAGTGGTTTCTACCGCAACAACAACTTTCTCGGGACCTGATCAAGCGC 3569
Db	651 GGAGAGTCAAAAAGTGTATCAAGTCCGACCAAGACTTTCCGAAGCAACCGATATAAGCGG 592
Qy	3570 GTGTGCTGAGTGTCTACCATGGCGCTGGCTCGAAGACCCTAGCGGTCCAAAAGGTCCA 3629
Db	591 GTGTGATGGACAGTATACAGGGGCGCGGAACAAGGACCATCGGTTCACCAAGGGTCAT 532
Qy	3630 ATCACCCAAATGTACACCAATGTAGACTGTGACCTCGTCGGCTGGCAGGCGCCCGCGG 3689
Db	531 GTTATCCAGATGTACACCAATGTGACCAAGCTTGTAGGCTGGCCCGCTCCCAAGGT 472
Qy	3690 CGCGGCTCCATGACACCATGAGTGTGGCAGCTCGGACCTTTACTTGGTCCACGAGCAT 3749
Db	471 GCTCGCTCATTAACACCTGTCACTTGGCGCTCTCGGACCTTACCTGCTCAGGAGCAC 412
Qy	3750 GCTGATGTATTCGGTGGCGCGGAGGAGCAGCAGGGAAGTCTACTCTCCCGCAGG 3809
Db	411 GCCGATGTCAATCCCGTGGCGGGGGGTGACAGCAGAGGCGAGCTGCTCTCGCGCGG 352
Qy	3810 CCCGTCCTTACCTAAAGGCTCTCGGGTGTGCTTTCCTTTCGCTTCGCGGACAGCTC 3869
Db	351 CCTATCTTCTTAAGGGCTCTCGGGGGGCGCCACTGCTGTGCCCCGCGGAGACCGCT 292
Qy	3870 GTGGCGGCTTCCTCGGGCTGTGTGTGCACCCGGGGGGTCCGGAAGCGGTGGACTTCATA 3929
Db	291 GTAGGCAATTTAGGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 232
Qy	3930 CCCGTTGAGTCTATGAAACTACCATGGGCTCTCGGCTCTTTCACAGACAACCTCAACCCC 3989
Db	231 CCCGTAGAAAGCCTAGAGCAACCATGAGTCCCGGGTTCACGAGCAACTCTCTCCCA 172
Qy	3990 CCGGCTGTACCGCAGACATTCOAAGTGGCAATCTGCACGGCTCTACTTGGCAGCGGCAAG 4049

Db	171 CCAGCAGTCCCCAGAGTTCAGGTGGCTCACCTGCATGCTCCCAACCGCAGCGTAAAG 112
Qy	4050 AGCACAAAAGTGCAGGCTCGTATGCAGCCCAAGGGTACAAGGTGCTCTTCTTGAACCCG 4109
Db	111 AGCACAAAGTCCCGCGCGGTACGGGCTCAGGGCTACAAGGTGGGGTGTCTCAACCCC 52
Qy	4110 TCGGTTCGCGCACCTTAGGTTTGGGGCGTATATGTCCTCAAGGCAC 4155
Db	51 TCGGTTGCTGCAACACTGGGCTTGGAGCGATAGAGGAATCGCGC 6

RESULT 4

CZ990230/c

LOCUS

DEFINITION

207284 Tomato MboI BAC Library Lycopersicon esculentum genomic clone SL_MboI0131A03 5, genomic survey sequence.

ACCESSION

CZ990230.1

VERSION

GI:72341875

KEYWORDS

GSS.

SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM

Lycopersicon esculentum

REFERENCE

1 (bases 1 to 769)
Mueller,L.A., Bueis,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J., Van Eck,J. and Stack,S.
BAC end sequencing from three Solanum lycopersicon libraries

AUTHORS

Unpublished (2005)

TITLE

BAC end sequencing from three Solanum lycopersicon libraries

JOURNAL

Cornell University

COMMENT

Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@sgn.cornell.edu
Plate: 131 row: A column: 3
Seq primer: T7
Class: BAC ends
High quality sequence start: 6
High quality sequence stop: 374.

FEATURES

source

1..769
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="SL_MboI0131A03"
 /lab_host="E. coli"
 /clone_lib="Tomato MboI BAC Library"
 /note="Vector: pBelOBAC11; Site_1: MboI"

ORIGIN

Query Match 3.3%; Score 318; DB 10; Length 769;
Best Local Similarity 74.0%; Pred. No. 7.2e-66;
Matches 416; Conservative 0; Mismatches 145; Indels 1; Gaps 1;

Qy	3551 GCGGACCTGCATCAACGGCGTGTCTGGACTGTCTTACATGGCGCTGGCTCGAAGACCT 3610
Db	562 GACAACCTTGACACAACAGGATGGGAGGAACGGTTACCCAGGGGCGAGAAATGGAGCTT 503
Qy	3611 AGCGGCTCCAAAAGGTCCAATCACCCA-AATGTACACCAATGTAGACTGACCTCGTCG 3669
Db	502 GGAGTCACACAAGGGACCTGTAGTCCAGAAATGTACCAATGTGGACCAAGACCTAGTG 443
Qy	3670 GCTGGCAGGCGCCCCCGGGCGGCTCCATGACACCATGACGCTGTGCGAGCTCGGACC 3729
Db	442 GCTGGCGGCTCCCGAGGGGCGCGCTCATTAACACCTGCACTTGGCGCTCTCGGACC 383
Qy	3730 TTTTACTTGGTCCAGAGACATGCTGATGTCACTTCGGTTCGCGCGGCGAGCGCAGCAGGG 3789

Db 382 TTTACTGTCACGAGCACGCCGATGTCATTCCTGCGCGCGGGTGACACGAG 323
 QY 3790 GAAAGTCTACTCTCCCGCCAGGCCGCTCTCTACTCTGAAAGGCTCTCTCGGGTGGTCATTGCG 3849
 Db 322 GCAGCTGCTCTCGCGCCGCGCTATCTCTTACTTGAAAGGCTCTCTCGGGGGCCCATGCG 263
 QY 3850 TTTGCCCTTCGGGGCACGTCGCTGGCGCTCTTCGCGGCTGCTGTGTGTCACCCCGGGGGTCG 3909
 Db 262 TGTGCCCCCGCGGACACGCTGTAGGCATATTAGGGCCGCGGTGTGCACCCGTGGAGTGG 203
 QY 3910 CGAAGCGGTGAGCTTCATACCCGTTGAGTGTATGGAATACATCGCGTCTCGGTCCT 3969
 Db 202 CTAAGCGGTGGAATTTGTCCCCGTAGAAAGCTTAGAGACAACCATGAGGTCCTCCCGGTGT 143
 QY 3970 TCACAGACAATCAACCCCCCGGCTGTACCGCACACATTCCAAGTGGCACATCTGCACG 4029
 Db 142 TCACGGAACATCTCTCCCAACAGCAGTGCCCGCAGCTTCCAGGTGGCTCACCTGCGATG 83
 QY 4030 CTCCTACTGCGCAGCGCACAGCACCAAGTGCCGGCTGCGTATGCAAGCCCAAGGGTACA 4089
 Db 82 CTCACCGCGCAGCGGTAAAGCACCAAGTCCCGCGCGGTACGCGGTACAGGGCTACA 23
 QY 4090 AGGTGCTGCTCTGAACCCGTC 4111
 Db 22 AGGTGCTGCTCAACCCCTC 1

RESULT 5
 CZ990303
 LOCUS
 DEFINITION 207384 Tomato MboI BAC Library Lycopersicon esculentum genomic
 clone SL_MboI0131C05 5, genomic survey sequence.
 ACCESSION CZ990303.1 GI:72341948
 KEYWORDS GSS.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 852)
 AUTHORS Mueller,L.A., Buel,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
 Van Eck,J. and Stack,S.
 TITLE BAC end sequencing from three Solanum lycopersicon libraries
 JOURNAL Unpublished (2005)
 COMMENT Other_GSSs: 207383
 Contact: Lukas Mueller
 Tanksley Lab, Dept. of Plant Breeding
 Cornell University
 251 Emerson Hall, Ithaca, NY 14853, USA
 Tel: 607-255-6557
 Fax: 607-255-6683
 Email: sgn-feedback@sgn.cornell.edu
 Plate: 131 row: C column: 5
 Seq primer: T7
 Class: BAC ends
 High quality sequence start: 28
 High quality sequence stop: 477.
 Location/Qualifiers
 1. .852
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="SL_MboI0131C05"
 /lab_host="E. coli"
 /clone_lib="Tomato MboI BAC Library"
 /note="Vector: pBelOBAC11; Site_1: MboI"

FEATURES
 source
 1. .852
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="SL_MboI0131C05"
 /lab_host="E. coli"
 /clone_lib="Tomato MboI BAC Library"
 /note="Vector: pBelOBAC11; Site_1: MboI"
 ORIGIN
 Query Match 3.3%; Score 315.2; DB 10; Length 852;
 Best Local Similarity 70.2%; Pred. No. 3.6e-65;
 Matches 438; Conservative 0; Mismatches 183; Indels 3; Gaps 1;

QY 3322 GGCACATCATCTTTGGGTCTACCCGTCTCCGCGCCGAGGGGAGAGATATTTTGGGAC 3381
 Db 23 GGGGAATCATCGACGGCTTGCCTCTCTGCCGTAGGGCCGAGAGATCTGCTCGGGC 82
 QY 3382 CGGCTGATAGTCTCGAAGGGCAAGGGTGGCGACTCTCTTGGCGCCCATCATCGGCTTACTCCC 3441
 Db 83 CGGCGATGAAATGGCTTCTAAGGGGTGGAGTTGTGGCGCCCATCACGGCGTATGCCC 142
 QY 3442 AACAAACGCGGGCGTACTTGGTTGCATCATCTAGCTACAGCCGCGGACAGAAC 3501
 Db 143 AGCAGACGAGGGGCTCTTGGGGTGCATAATTACAGCCTGACCGCGGACAGAAC 202
 QY 3502 AGGTCTGAAGGGGAGGTTCAAGTGGTTCTACCGCAACACAATCTTCTCGCGACCTGCA 3561
 Db 203 AGGTGGAGGGTGAAGTTGAGTTGTGTCACTGCGGCCAGACCTTCTTGGCACTGTA 262
 QY 3562 TCACACGGGTGTGTGACTGTCTACCATGGCGTGGCTCGAAGACCTAGCCGGTCCAA 3621
 Db 263 TTAACGGGTGTGTGACCGTCTACCAAGGGGCGGAACAAGGACCATCGCGTCACCCA 322
 QY 3622 AAGTCCAATCACCAATGTACACCAATGTAGACCTGGACCTCGTGGCTGCGGCGGCGC 3681
 Db 323 AGGTCTCTGTATTCCAGATGTACCAATGTGGACCAAGACCTGTAGGCTGGCGCCGCTC 382
 QY 3682 CCCCAGGGGCGGCTCCATGACACCATGTCAGCTGTGGCAGCTCGGACCTTTACTTTGGTCA 3741
 Db 383 CCCAAGGTGCGGCTCATTAACACCTGCACTTGGGCTCTCTCGGACCTTTACCTGGTCA 442
 QY 3742 CGACATATGTGATGTCTATCCGTCGCGCGGCGAGCGACAGCAGGGGAGTCTACTCT 3801
 Db 443 CGAGGCACGCGATGTCTATCCCGTGGCGCGGGGTGACATCATAGGCATCTCTGCTCT 502
 QY 3802 CCCCAGGGGCGGCTCTCTACTCTGAAAGGCTCTCGGTGGTTCATTGTTGCCCTTCGG 3861
 Db 503 CGCCCGGCTTATCTCTTACTTGAAGGCTCTCTCGGGGGGCCACTTCTTGGTCCCCTG 562
 QY 3862 GG---CAGCTGCTGGCGCTCTTCGGGCTGTGTGTGCACCGGGGGTCCGGAAGCGG 3918
 Db 563 GGATACCTGTACGTATATTATAGCCCGTGTGTGCTCCCTGTTATGCGCTTAGGCC 622
 QY 3919 TGGACTTTCATCCCGTTGAGTCTA 3942
 Db 623 TTGTATTTCATCCCTTCAGACTA 646

RESULT 6
 CZ990242
 LOCUS
 DEFINITION 207300 Tomato MboI BAC Library Lycopersicon esculentum genomic
 clone SL_MboI0131A11 5, genomic survey sequence.
 ACCESSION CZ990242
 VERSION CZ990242.1 GI:72341897
 KEYWORDS GSS.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 817)
 AUTHORS Mueller,L.A., Buel,R.M., Wang,Y., Tanksley,S.D., Giovannoni,J.J.,
 Van Eck,J. and Stack,S.
 TITLE BAC end sequencing from three Solanum lycopersicon libraries
 JOURNAL Unpublished (2005)
 COMMENT Other_GSSs: 207299
 Contact: Lukas Mueller
 Tanksley Lab, Dept. of Plant Breeding
 Cornell University
 251 Emerson Hall, Ithaca, NY 14853, USA
 Tel: 607-255-6557
 Fax: 607-255-6683
 Email: sgn-feedback@sgn.cornell.edu
 Plate: 131 row: A column: 11

```
Seq primer: T7
Class: BAC ends
High quality sequence start: 5
High quality sequence stop: 420.
FEATURES
    source
        1..817
            /organism="Lycopersicon esculentum"
            /mol_type="genomic DNA"
            /cultivar="Heinz 1706"
            /db_xref="taxon:4081"
            /clone="SL MboI0131A11"
            /lab_host="E. coli"
            /clone_lib="Tomato MboI BAC Library"
            /notes="Vector: pBelOBAC11; Site_1: MboI"
ORIGIN
    Query Match      3.0%; Score 287.2; DB 10; Length 817;
    Best Local Similarity 70.7%; Pred. No. 2.5e-58;
    Matches 382; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
QY 3340 TACCCGTCCTCCGCCGGAAGGGGAGGAGATATTTTGGGACCGGCTGTATAGTCTCGAAG 3399
Db 4 TGGCCGTCCTCGCCGTAGGGCGGAGAGATCTGCTCGGGCGCGCGATGGATGGCTT 63
QY 3400 GGCAAGGGTGGCGACTCTTTCGCCCATCAAGGCTTACTCCCAACAACGCGGGCGGTAC 3459
Db 64 CTAAGGGGTGGAGTTGTTCGGCGCCATCATCGGCGTATGCCAGCAGCAGAGAGCGCTCT 123
QY 3460 TTGGTTGCATCATCACTAGCGCTCAAGCGCGGGAAGAAGACAGGTTCGAAGGGGAGTTTC 3519
Db 124 TGGGGTGCATAATATACAGGCTTGACCGCGCGGGAAGAAGACAGGTTCGAGGTTTC 183
QY 3520 AAGTGGTTTCTACCGCAACATCTTCTCGCGACTCGATCAACAGGGGTGTCGGA 3579
Db 184 AGATTGTGTAACCTGCGCGCGAGACCTTCTCGCAACCTGTATTAACAGGGGTGTCGGA 243
QY 3580 CTGCTACCATGCGGCTGGTTCGAAGACCTCTAGCGGTCGCAAAAGTCCATCAACCAAA 3639
Db 244 CCGTCTACCAAGCGGCGGACAGGACCATCGGTCACCAAGGGTCTGTATCCAGA 303
QY 3640 TGTAACCAATGTAGACTCGTTCGGCTGGCAGGGCGCCCGCGGGCGCGCTCCA 3699
Db 304 TGTACACCAATGTGGAACCAAGACTCTGAGCTGGCGCGCTCCCAAGGTCCCGCTCAT 363
QY 3700 TGACACCATGAGCTTGGAGCTCGGACCTTTTCTGTTGTCACAGACATGCTCATGTCA 3759
Db 364 TAACACCTGTCACCTTGGGGTCTCTCGACCTTTTACCTGGTCACAGGACACCGCATGTCA 423
QY 3760 TTCCGGTGGCGCGGCGGAGGACAGCGGGAAGTCTACTCTCCCGCAGGCGCGCTCTCCT 3819
Db 424 TTCCCGTGGCGCGGCGGAGTACAGAGGACGACGCTGCTCTCCCTCGCGCTATCTCTT 483
QY 3820 ACCTGAAAGGCTCTCGGGTGGTCAATTCGCTTTTGGCCCTTTCGGGGCAGCTCGTGGGCGTCT 3879
Db 484 ACTTGAGGGGCTCTCTCGGGGCGCCCTGCTCTAGCCCTTCAGGACATCCGCTGTAT 543
RESULT 7
CZ990744 922 bp DNA linear GSS 11-AUG-2005
LOCUS 207968 Tomato MboI BAC Library Lycopersicon esculentum genomic
DEFINITION clone SL MboI0131009 5, genomic survey sequence.
ACCESSION CZ990744.1 GI:72342389
VERSION 1
KEYWORDS GSS.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 922)
REFERENCE
AUTHORS Mueller, L.A., Buels, R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,
van Eck, J. and Stack, S.
BAC end sequencing from three Solanum lycopersicon libraries
Unpublished (2005)
Other_GSSs: 207967
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@sgn.cornell.edu
Plate: 131 row: O column: 9
Seq primer: T7
Class: BAC ends
High quality sequence start: 48
High quality sequence stop: 274.
FEATURES
    source
        1..922
            /organism="Lycopersicon esculentum"
            /mol_type="genomic DNA"
            /cultivar="Heinz 1706"
            /db_xref="taxon:4081"
            /clone="SL MboI0131009"
            /lab_host="E. coli"
            /clone_lib="Tomato MboI BAC Library"
            /notes="Vector: pBelOBAC11; Site_1: MboI"
ORIGIN
    Query Match      2.7%; Score 257.4; DB 10; Length 922;
    Best Local Similarity 84.6%; Pred. No. 4.8e-51;
    Matches 324; Conservative 0; Mismatches 56; Indels 3; Gaps 3;
QY 3315 GCGTGTGGGACATCATCTTTGGGTCTACCGCTCTCCGCCCAAGGGGGAAGAGATATT 3374
Db 5 GCGTGTGGGACAT-ATCGTGGGCTGCGCTCTCCGCCCGCAGGGGGAAGAGATATT 63
QY 3375 TTGGGACCGGCTGATAGTCTCGAAGGCAAGGTGGGAGCTCTTGGCGCCCATCACGGCC 3434
Db 64 CTGGCGCGGCTGATAGCTTGAAGGACAGGGGTGGCGGCTCTTGGCGCTATCACGGCT 123
QY 3435 TACTCCCAACAACGCGGGCGGTACTTTGGTTCATCATCTAGCCTCACAGGCGGGAC 3494
Db 124 TACTCCAGCAGACGCGGGGCTTACTTTGGTTCATCATCTAGCCTCACAGGCGGGAC 183
QY 3495 AAGAACACAGGTTCGAAGGGGAGGTTCAAGTGTCTTACCGCAACAATCTTCTCGCG 3554
Db 184 AAGAACACAGTTCGAGGGGAGGTTCAAGTGTCTTACCGCGCACAGTCTTCTCGCG 243
QY 3555 ACCTGCATCAACGCGGTGCTGAGCTGTCTACCATGGCGTGGCTCGAAGACCTAGCC 3614
Db 244 ACCTGCATCAATGGCGGTGTGTGGACTGTCTACCATGGCGCGCTCAAGACCTAGCC 303
QY 3615 GTTCCAAAAGTCCAATCA-CCCAATGTACCAATGTA-GACCTGGACCTCGTCGGCT 3672
Db 304 GGCCGAATGGCCCAATCATCCCAAGTTGTACCAAGTGTATACCAAGATCTCTCGGAT 363
QY 3673 GGCAGGCGCGCCCGGGCGCGC 3695
Db 364 TGTCAACCGCTCTGAGGAC 386
RESULT 8
CZ990535 891 bp DNA linear GSS 11-AUG-2005
LOCUS 207704 Tomato MboI BAC Library Lycopersicon esculentum genomic
DEFINITION clone SL MboI0131121 5, genomic survey sequence.
ACCESSION CZ990535
VERSION 1
KEYWORDS GSS.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
```

REFERENCE 1 (bases 1 to 891)
 AUTHORS Mueller,L.A., Buels,R.M., Wang,Y., Tankley,S.D., Giovannoni,J.J.,
 Van Eck,J. and Stack,S.
 TITLE BAC end sequencing from three Solanum lycopersicon libraries
 JOURNAL Unpublished (2005)
 COMMENT Other_GSSs: 207703
 Contact: Lukas Mueller
 Tanksley Lab, Dept. of Plant Breeding
 Cornell University
 251 Emerson Hall, Ithaca, NY 14853, USA
 Tel: 607-255-6557
 Fax: 607-255-6683
 Email: sgn-feedback@cornell.edu
 Plate: 131 row: I column: 21
 Seq primer: T7
 Class: BAC ends
 High quality sequence start: 39
 High quality sequence stop: 261.

FEATURES
 source
 1..891
 Location/Qualifiers
 /organism="Solanum lycopersicon"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="SL_MboI131121"
 /lab_host="E. coli"
 /clone_lib="Tomato MboI BAC Library"
 /note="Vector: pBelBAC11; Site_1: MboI"

ORIGIN
 Query Match 1.5%; Score 139.6; DB 10; Length 891;
 Best Local Similarity 71.3%; Pred. No. 2.9e-22;
 Matches 184; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 3349 CCGCCCGAGGGGAGAGATATTTTGGACCGCTGATGTCGAGGGCAAGGT 3408
 DB 30 CTGACGGTAGGGCCGAGAGATACTGCTCGGCCCGCGATGGAATGGCTTTAAGGGT 89

QY 3409 GCGCACTCTTCGCGCCATCAGCGCTACTCCAAACAAACGGGGCGTACTTGGTGCA 3468
 DB 90 GGAGTTGTTGGCGCCATCAGCGGTATGCCAGACAGAGGGGCTCTTGGGGTCA 149

QY 3469 TCATCACTAGCTCAGCGCGGCAAGAACAGGTGCAAGGGAGGTTCAAGTGGTTT 3528
 DB 150 TAAATTACAGCTCAGCGCGGCAAGAACAGGTGAGGGTTCATATTTGT 209

QY 3529 CTACGCGACAAATCTTCTGGGACCTGATCAACGGGTGCTGGACTGTCTACC 3588
 DB 210 CAACTGCGGCCAGACCTTCTCGCAACCTGATTAACGGGTGTGCTGGACCGTCTACC 269

QY 3589 ATGGCGCTGGCTCGAAGA 3606
 DB 270 ACGGGCGCGAACAGGA 287

RESULT 9
 AV755731/c 487 bp mRNA linear EST 22-NOV-2004
 LOCUS AV755731 BM Homo sapiens cDNA clone BFAKB03 5', mRNA sequence.
 DEFINITION AV755731
 ACCESSION AV755731
 VERSION AV755731.2 GI:55952773
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.
 1 (bases 1 to 487)
 Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
 Gu,Y., Li,N., Qian,B., Liu,F., Ou,J., Gao,X., Cheng,Z., Xu,Z.,
 Zeng,L., Xu,S., Gu,Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,W.,
 Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and
 Chen,Z.

REFERENCE 10
 CV350260/c 597 bp mRNA linear EST 27-SEP-2004
 LOCUS CV350260
 DEFINITION MR2-SN0006-050600-002-b11 SN0006 Homo sapiens cDNA, mRNA sequence.
 ACCESSION CV350260
 VERSION CV350260.1 GI:52700315
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.

TITLE Homo sapiens cDNA BM clones
 JOURNAL Unpublished (2000)
 COMMENT On Oct 19, 2000 this sequence version replaced gi:10913579.
 Contact: Zengqiang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzq@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES
 source
 1..487
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BFAKB03"
 /tissue_type="Bone marrow"
 /cell_type="CD34+ hematopoietic stem/progenitor cell"
 /lab_host="BM25.8"
 /clone_lib="BM"
 /note="Vector: pTriplEx2; Site_1: sf1A; Site_2: sf1B"

ORIGIN
 Query Match 1.3%; Score 129; DB 1; Length 487;
 Best Local Similarity 58.9%; Pred. No. 9.5e-20;
 Matches 275; Conservative 0; Mismatches 187; Indels 5; Gaps 3;

QY 1137 GTTGGGACGGCTGCTTCTGCTCGCTATGTACGTGGGGATCTCTGGGATCTATTTC 1196
 DB 472 GTGGTGTACACCTCGCTCTGCTCAGCTCTTACGTGGGACCTCTGCGAGGATGATG 413

QY 1197 CTCGTCTCCAGCTGTTACACCTCTCGCTCGCGGATGAGACAGTCAGGACTGCAAC 1256
 DB 412 CTTGAGTTGAGCTGATCA---TCTGGCTCAGCACCATGAGTTTGTGCATGAATGCAAC 356

QY 1257 TGCTCAATATATCCCGGCATGTATCAGGTACCGCATG-CGTTGGGATATGATGAA 1315
 DB 355 TGCTCCATCTATCTGCGGCCATCACTGGACACCGTATGAGCATGGGACATGATGAA 296

QY 1316 CTGTCACCTACACAGCCCTAGTGGTGTGCGAGTTGTCGGGATCCCAAGCTGTCTG 1375
 DB 295 CTGTCGTGACCCCTGCTATGATCATGGCGTACGCAATGCGCGTCTCTGAGGTCATCAT 236

QY 1376 GGACATGTTGGCGGGGCCACTGGGAGTCTCGCGGCGCTTGCTACTATTCCATGCT 1435
 DB 235 AGATATCATCAGCGGGGCTCACTGGGGGCTCATGTTGGGCTTAGGCTACTTCTCTATGCA 176

QY 1436 AGGAACTGGGCTAAGGTTCTGATTGTGGCGCTACTCTTTGCCGCGTGTGACGGGAGAC 1495
 DB 175 GGGAGCGTGGCGAAGTCTGCTCTCTGTGACCGCTGGGGTACACGCGAGCAC 116

QY 1496 CCACAGCGGGGAGGTGGCGGCG-CACACACCTCGGGTTACGTCCTTTTCTCAT 1554
 DB 115 CCAGTCCACAGTGGCCTGCTAGCGCGCGCCAGCAGTTTCGTCAAGTTGTCTCAACC 56

QY 1555 CTGGGCGCTCTCAGAAATCCAGCTTGTGAATACCAACGCGAGCTGG 1601
 DB 55 CTGGCGCTAAACAGAACATCCAACCTATTANGATCNATTGCAATGGG 9

RESULT 10
 CV350260/c 597 bp mRNA linear EST 27-SEP-2004
 LOCUS CV350260
 DEFINITION MR2-SN0006-050600-002-b11 SN0006 Homo sapiens cDNA, mRNA sequence.
 ACCESSION CV350260
 VERSION CV350260.1 GI:52700315
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.

REFERENCE 1 (bases 1 to 597)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. <http://www.ludwig.org.br>.

FEATURES
source
1. 597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0006"
/notes="Organ: stomach_normal; Vector: puc18; Site_1: Smal; Site_2: Smal; A mini-library was made by cloning products derived from ORFESTES PCR (U.S. Letters Patent application No. 196, 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 1.1%; Score 102.8; DB 7; Length 597;
Best Local Similarity 86.9%; Pred. No. 2.5e-13; Indels 0; Gaps 0;
Matches 113; Conservative 0; Mismatches 17;
QY 184 GACCGGTCCTTTCTTGATCAACCGCTCAATGCTGGAGATTGGCGGTGCCCGCG 243
Db |||||
130 GACCGGTCCTTTCTTGGAACACCGCTCAATACCCAGAAATTGGCGGTGCCCGCG 71
QY 244 AGATGCTAGCCGAGTAGTGTGGGTGCGAAAGCCCTTGTGTGCTGCTGTAGAGTG 303
Db |||||
70 AGATCCCTAGCCGAGTAGTGTGGGTGCGAAAGCCCTTGTGTGCTGCTGTAGAGG 11
QY 304 CTTGGGAGTG 313
Db |||||
10 TCCTGGAGG 1

RESULT 11
AV758366/c
LOCUS AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA linear EST 19-OCT-2000
DEFINITION AV758366
ACCESSION AV758366
VERSION AV758366.1 GI:10916214
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 492)
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
TITLE Homo sapiens cDNA BM clones

JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1. 492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB"

ORIGIN
Query Match 1.0%; Score 98; DB 1; Length 492;
Best Local Similarity 55.2%; Pred. No. 3.6e-12; Indels 5; Gaps 3;
Matches 250; Conservative 0; Mismatches 198;
QY 1153 TCTGCTCCGCTATGTAGTGGGGATCTCTGCGGATCTATTTTCCTGCTCCAGCTGT 1212
Db |||||
457 TGTGATCAGCTCACTACGTTGCGCTCTGCGTTGGGGTATCGTTGCAGCCCACTGA 398
QY 1213 TCACCTTTCGCTCGCGGCATGACAGCTGAGAGCTGCAACTGCTCACTATATCCCG 1272
Db |||||
397 TTA---TCTCTCAGCAGCAACATTGGTTGTGCAAGAATGCAACTGCTCACTATCTCTG 341
QY 1273 GCCATGTATCAGTGC-ACGCGATGGCTTGGGATATGATGATGAATGCTCACTTACAACA 1331
Db |||||
340 GCTGCATCACTGGACTACAGTATGGATAGCTATGATGATGAATGCTGGGACCGGT 281
QY 1332 GCCTAGTGTGTGCGAGTTGCTCCGGATCCCAAGCTGCTGGGACATGTTGGCGGG 1391
Db |||||
280 TCCATGATATGCGGTACGCAATGCGGTTCTCTGAAGTCTCATAGATATCATAAGCTGG 221
QY 1392 GCCACTGGGAGTCTTGGCGGCTTGCCTACTATTTCCATGTTAGGAACTGGGCTAAG 1451
Db |||||
220 GCACACTGCGCGCTGATGTTGCGCTCAGCTTACTTCAATGCAAGGAGCGTTGGCCAAA 161
QY 1452 GTTCTGATTTGGCGGTACTCTTTTGGCGGTTGACGGGAGACCCACACAGCGGAGG 1511
Db |||||
160 GTAGCTGTATCCATCTGTTTCACGCTGGGGTTGATCGGAGCAACACGATATCAGGTGGT 101
QY 1512 GTGGCGG-GCCACACCACTCCGGGTTTCAGTCTCTTTTCTCATCTGGGGCGTCTCAGAA 1570
Db |||||
100 ATCACTGCGCGGACGACAGAGTTTCGTAGTTGATTAATGCTGCGCTAGCAGTA 41
QY 1571 AATCCAGCTTTGTGAATACCAACCGCAGCTGSCA 1603
Db |||||
40 CATCCAGCTCATGANCCTCAATGSCAATNGNA 8

RESULT 12
CN648295/c
LOCUS CN648295 832 bp mRNA linear EST 13-MAY-2004
DEFINITION ILLUMIGEN MQ0 30183 Katze MMPB Macaca mulatta cDNA clone IBUM.6735 5' similar to Bases 150 to 481 highly similar to human FLJ14743 (Hs.169577), mRNA sequence.
ACCESSION CN648295
VERSION CN648295.1 GI:47161738
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopitheidae; Cercopithecinae; Macaca.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2005, 08:15:42 ; Search time 418 Seconds
(without alignments)
5080.479 Million cell updates/sec

Title: US-09-662-454-3

Perfect score: 16009

Sequence: 1 MSTNPKPQKTKRNTNRFP.....FPLCLLLSVGVGYLLPNR 3010

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16009	100.0	3010	2	O92969_9HEPC
2	15949	99.6	3010	2	O92970_9HEPC
3	15937	99.6	3010	2	O92971_9HEPC
4	15909	99.4	3010	2	O92972_9HEPC
5	15887	99.2	3010	2	O92973_9HEPC
6	15772	98.5	3010	2	O92974_9HEPC
7	15722	98.3	3010	2	O92975_9HEPC
8	15392	96.1	3010	2	O92976_9HEPC
9	15381	96.1	3010	2	O92977_9HEPC
10	15378	96.1	3010	2	O92978_9HEPC
11	15362	96.0	3009	1	POLG_HCVJA
12	15358	95.9	3010	2	O92979_9HEPC
13	15356	95.9	3010	2	O92980_9HEPC
14	15338	95.8	3010	2	O92981_9HEPC
15	15330	95.8	3010	2	O92982_9HEPC
16	15326	95.7	3010	2	O92983_9HEPC
17	15323	95.7	3010	2	O92984_9HEPC
18	15322	95.7	3010	2	O92985_9HEPC
19	15319	95.7	3010	2	O92986_9HEPC
20	15318	95.7	3009	1	POLG_HCVCO
21	15303	95.6	3010	2	O92987_9HEPC
22	15301	95.6	3010	2	O92988_9HEPC
23	15299	95.6	3010	2	O92989_9HEPC
24	15299	95.6	3010	2	O92990_9HEPC
25	15280	95.4	3010	2	O92991_9HEPC
26	15278	95.4	3010	2	O92992_9HEPC
27	15277	95.4	3010	2	O92993_9HEPC
28	15273	95.4	3010	2	O92994_9HEPC
29	15272	95.4	3010	2	O92995_9HEPC
30	15272	95.4	3010	2	O92996_9HEPC
31	15267	95.4	3010	2	O92997_9HEPC

32	15267	95.4	3010	2	O5R2C5_9HEPC	O5R2C5 hepatitis c
33	15267	95.4	3010	2	O9J3G6_9HEPC	O9J3G6 hepatitis c
34	15265	95.4	3010	2	O8I760_9HEPC	O8I760 hepatitis c
35	15263	95.3	3010	2	O5R2B5_9HEPC	O5R2B5 hepatitis c
36	15253	95.3	3010	2	O9I3V3_9HEPC	O9I3V3 hepatitis c
37	15253	95.3	3010	2	O5R2B1_9HEPC	O5R2B1 hepatitis c
38	15253	95.3	3010	2	O8V638_9HEPC	O8V638 hepatitis c
39	15250	95.3	3010	2	O5R2B2_9HEPC	O5R2B2 hepatitis c
40	15248	95.2	3009	1	POLG_HCVBK	P26663 h genome po
41	15246	95.2	3010	2	O5R2B0_9HEPC	O5R2B0 hepatitis c
42	15245	95.2	3010	2	O5R2D4_9HEPC	O5R2D4 hepatitis c
43	15243	95.2	3010	2	O5R2C1_9HEPC	O5R2C1 hepatitis c
44	15242	95.2	3010	2	O9Q1Y5_9HEPC	O9Q1Y5 hepatitis c
45	15241	95.2	3010	2	O5R2C4_9HEPC	O5R2C4 hepatitis c

ALIGNMENTS

RESULT 1
ID O92969_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O92969;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo."
RL Virology 244:161-172(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity."
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF054247; AAC15722.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; P02446; P02446.
DR PIR; P80329; P80329.
DR HSSP; Q02828; 1NB4.
DR SMR; Q92969; 1029-1657, 2008-2170, 2420-2949.
DR GO; Q0019028; C:viral capsid; IEA.
DR GO; Q0019031; C:viral envelope; IEA.
DR GO; Q0005524; F:ATP binding; IEA.
DR GO; Q0008026; F:ATP-dependent helicase activity; IEA.
DR GO; Q0003723; F:RNA binding; IEA.
DR GO; Q0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; Q0008236; F:serine-type peptidase activity; IEA.
DR GO; Q0005198; F:structural molecule activity; IEA.
DR GO; Q0006508; P:proteolysis and peptidolysis; IEA.
DR GO; Q0006350; P:transcription; IEA.
DR GO; Q0019079; P:viral genome replication; IEA.
DR GO; Q0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.

DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRP.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326781 MW; 9B3FD910CF00E2C5 CRC64;

Query Match 100.0%; Score 16009; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKASERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKASERSQPRG 60
QY 61 RRQPIKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPDPRRRSRLG 120
DB 61 RRQPIKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPDPRRRSRLG 120
QY 121 KVIDLTCGFADLMGYIPLVGAPLGAARALAHGVRLVEDGVNYATGNLPGCSPSIFLLA 180
DB 121 KVIDLTCGFADLMGYIPLVGAPLGAARALAHGVRLVEDGVNYATGNLPGCSPSIFLLA 180
QY 181 LLSCLTIPASAEYVRNVSGIYHVTNDCSNLSIVYEAAADVIMHTPGCPVCQVQSGNSRCWV 240
DB 181 LLSCLTIPASAEYVRNVSGIYHVTNDCSNLSIVYEAAADVIMHTPGCPVCQVQSGNSRCWV 240
QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVGDLGSIIFLVSLFTFSRRRHET 300
DB 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVGDLGSIIFLVSLFTFSRRRHET 300
QY 301 VQDCNCSIYPGHVSGHRMAWDMNMWSPPTALVVSQLLRIPQAVDMVAGAHGWVLAGLA 360
DB 301 VQDCNCSIYPGHVSGHRMAWDMNMWSPPTALVVSQLLRIPQAVDMVAGAHGWVLAGLA 360
QY 361 YYSVMGNWAKVLIVALLFAGVDEGTHTTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
DB 361 YYSVMGNWAKVLIVALLFAGVDEGTHTTGRVAGHTTSGFTSLFSSGASOKIQLVNTNGSW 420
QY 421 HINRTALNCNDSIQTGFFAALFYAHKFNSGCGPERMASCRPIDWFAQGWGPITYTKPNSS 480
DB 421 HINRTALNCNDSIQTGFFAALFYAHKFNSGCGPERMASCRPIDWFAQGWGPITYTKPNSS 480
QY 481 DQRPYCHWHAPRCGVVPASQVCGPVYCFPTSPVVVGTTRDGSVPYTSWGENETDVMLLN 540
DB 481 DQRPYCHWHAPRCGVVPASQVCGPVYCFPTSPVVVGTTRDGSVPYTSWGENETDVMLLN 540
QY 541 NTRPPQGNWFGCTWMNSTGFTKCGGPPCNIIGVGNRTLICTDCPRKHPEATYTKCGSG 600
DB 541 NTRPPQGNWFGCTWMNSTGFTKCGGPPCNIIGVGNRTLICTDCPRKHPEATYTKCGSG 600
QY 601 PMLTPRCLVDYPYRLWHYPCTLNFSIFKVRMYVGGVHEHRLNAACNWRGERCNLEDRDS 660
DB 601 PMLTPRCLVDYPYRLWHYPCTLNFSIFKVRMYVGGVHEHRLNAACNWRGERCNLEDRDS 660
QY 661 ELSPLLLSTTEWQILPCAFPTTIPALSTGLIHQHONIVDVQVLYGVGSAFVSPFAIKWEYIL 720
DB 661 ELSPLLLSTTEWQILPCAFPTTIPALSTGLIHQHONIVDVQVLYGVGSAFVSPFAIKWEYIL 720

QY 721 LLFLLLLADARVCACLWMMLLIAQEAALLENLVLNAASVAGAHGILSLFVFFCAAWYIKG 780
DB 721 LLFLLLLADARVCACLWMMLLIAQEAALLENLVLNAASVAGAHGILSLFVFFCAAWYIKG 780
QY 781 RLAPGAAYAFYGVWPLLLLLLALPPRAYALDREMAASCGGAVLVGLVFLTISPYKVFIL 840
DB 781 RLAPGAAYAFYGVWPLLLLLLALPPRAYALDREMAASCGGAVLVGLVFLTISPYKVFIL 840
QY 841 RLIIWMLQYFTRABAHMQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGLPLM 900
DB 841 RLIIWMLQYFTRABAHMQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGLPLM 900
QY 901 VLQAGITRVPYFVRAOGLIRACMLVRKVAGGHYQVMFMKLGALTGTYYNHLTPLRDA 960
DB 901 VLQAGITRVPYFVRAOGLIRACMLVRKVAGGHYQVMFMKLGALTGTYYNHLTPLRDA 960
QY 961 HAGLRDLAVAVEPVVFSMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
DB 961 HAGLRDLAVAVEPVVFSMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
QY 1021 QGWRLLAPITAYSOQTRGVLCIITSLTGRDNQVGEVQVVSTATQSFATCINGVCWT 1080
DB 1021 QGWRLLAPITAYSOQTRGVLCIITSLTGRDNQVGEVQVVSTATQSFATCINGVCWT 1080
QY 1081 VYHAGSKTLAGPKGPIQMTYTNVDLVLVWQAPPGARSMTPCSCGSSDLVLTVRHADVI 1140
DB 1081 VYHAGSKTLAGPKGPIQMTYTNVDLVLVWQAPPGARSMTPCSCGSSDLVLTVRHADVI 1140
QY 1141 PVRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDPIPVES 1200
DB 1141 PVRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDPIPVES 1200
QY 1201 METTMRSPVFTDNSTPPPAVPQTFQVAHLHAPTGSKSTKVPAAYAAQGVKVLVLPNSVAA 1260
DB 1201 METTMRSPVFTDNSTPPPAVPQTFQVAHLHAPTGSKSTKVPAAYAAQGVKVLVLPNSVAA 1260
QY 1261 TLGFGAYMSKAHGIDPNIRTGVRTITTCGSIITYSGYKFLADGGSGGAYDIIICDECHS 1320
DB 1261 TLGFGAYMSKAHGIDPNIRTGVRTITTCGSIITYSGYKFLADGGSGGAYDIIICDECHS 1320
QY 1321 TDSITLIGITVLDQAETAGARLVVLTATPPGVSVPVPHNIEETGLSNNGEIPFYGKAI 1380
DB 1321 TDSITLIGITVLDQAETAGARLVVLTATPPGVSVPVPHNIEETGLSNNGEIPFYGKAI 1380
QY 1381 PIEAIKGRHLIFCHSKKKDELAAKLTCGLNAVAYVYRGDVSVPITGDVVVATDAL 1440
DB 1381 PIEAIKGRHLIFCHSKKKDELAAKLTCGLNAVAYVYRGDVSVPITGDVVVATDAL 1440
QY 1441 MTGFTGDPDSVIDCNTCVTQVDFSLDPTFTIETTVPQDAVRSQRRGRTGRSGIYR 1500
DB 1441 MTGFTGDPDSVIDCNTCVTQVDFSLDPTFTIETTVPQDAVRSQRRGRTGRSGIYR 1500
QY 1501 FVTPGERSGMPDSSVLCCEYDAGCAWYELTPAETSRLRAYLNTPTGLPVCODHLEFWES 1560
DB 1501 FVTPGERSGMPDSSVLCCEYDAGCAWYELTPAETSRLRAYLNTPTGLPVCODHLEFWES 1560
QY 1561 VPTGLTHIDAHFLSOTKQAGNFPYLVAQATVCARAQAPPSWDMKCLRLRKLPTLHG 1620
DB 1561 VPTGLTHIDAHFLSOTKQAGNFPYLVAQATVCARAQAPPSWDMKCLRLRKLPTLHG 1620
QY 1621 PTPLLYRLGAVQNEVILTHPITKYIMACMSADLEVVVTSTVLVGGVLAALAYCLTTGVS 1680
DB 1621 PTPLLYRLGAVQNEVILTHPITKYIMACMSADLEVVVTSTVLVGGVLAALAYCLTTGVS 1680
QY 1681 VIVGRILISGKPAVVPDREVLYQEFDEMEBCASQLPYIEQGMQLAEQFKQKALGLLQAT 1740
DB 1681 VIVGRILISGKPAVVPDREVLYQEFDEMEBCASQLPYIEQGMQLAEQFKQKALGLLQAT 1740
QY 1741 KQAEAAA PVESKWPALETFNAKHMWNFI SGIOYLAGLSTLPGNPAIASMAFTASITSP 1800
DB 1741 KQAEAAA PVESKWPALETFNAKHMWNFI SGIOYLAGLSTLPGNPAIASMAFTASITSP 1800
QY 1801 LTTQNTLLFNILGGWVAAQLAPPSSAASAFVGVAGIAGAAVGSIGLGKVLVDIILAGYGAGVA 1860

1801 LTTQNTLLFNILGSGWAAQLAPPASASAFVAGAGIAGAAGVSGIGLVLDILAGYGAGVA 1860
1861 GALVAFKWSGEVPESTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVQVWNRLLI 1920
1861 GALVAFKWSGEVPESTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVQVWNRLLI 1920
1921 AFASRGNHVSPTHYVPESDAAAARVTOILSSLTITQLLKRLHWINEDCSTPCSGSWLRDV 1980
1921 AFASRGNHVSPTHYVPESDAAAARVTOILSSLTITQLLKRLHWINEDCSTPCSGSWLRDV 1980
1981 WDWICTVLTDFKTLWOSKLLPRLPGVPFLSCQGYKGVWREGDGMOTTCPCGAQIAGHVK 2040
1981 WDWICTVLTDFKTLWOSKLLPRLPGVPFLSCQGYKGVWREGDGMOTTCPCGAQIAGHVK 2040
2041 NGSMRIVGPRCTCNTHGTPPINAYTTGCTPSPAPNYSRALNRVAEEVEVTRVGDFF 2100
2041 NGSMRIVGPRCTCNTHGTPPINAYTTGCTPSPAPNYSRALNRVAEEVEVTRVGDFF 2100
2101 YVTGMTDNDVKCQVPAPPEFFTEVDGVRLLHRVAPACKLLREDVTFQVGLNQYLVGSOL 2160
2101 YVTGMTDNDVKCQVPAPPEFFTEVDGVRLLHRVAPACKLLREDVTFQVGLNQYLVGSOL 2160
2161 PCPEPEPDVTLTSMLTDPSSHITAEAKRLARGSPPSLASSASOLSAPSKATCTTHHD 2220
2161 PCPEPEPDVTLTSMLTDPSSHITAEAKRLARGSPPSLASSASOLSAPSKATCTTHHD 2220
2221 SPADLIIEANLLWRQEMGNITRVESENKVVILDSFEPLHAEGDEREISVAABILKRSRK 2280
2221 SPADLIIEANLLWRQEMGNITRVESENKVVILDSFEPLHAEGDEREISVAABILKRSRK 2280
2281 FPSALPIWARPDYNNPPLLESKWOPDYVPPVHVGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
2281 FPSALPIWARPDYNNPPLLESKWOPDYVPPVHVGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
2341 SSALAEIATTFGSSGSANDSGTATLPDASDGDGKSDVESYSMPLEGEPPDIL 2400
2341 SSALAEIATTFGSSGSANDSGTATLPDASDGDGKSDVESYSMPLEGEPPDIL 2400
2401 SDGSWSVSEASDDVCCSMTYTWTCALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
2401 SDGSWSVSEASDDVCCSMTYTWTCALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
2461 SRASLAKQKVTFDRLQVLDHRYDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
2461 SRASLAKQKVTFDRLQVLDHRYDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
2521 GYGAKOVRNLSRAVNHRSVWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
2521 GYGAKOVRNLSRAVNHRSVWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
2581 FPDGLGVRCEKVALYDVVSTLPOAVMGSSYGFOYSQKQVBEFLVNTWKSCKCPMGFSYDT 2640
2581 FPDGLGVRCEKVALYDVVSTLPOAVMGSSYGFOYSQKQVBEFLVNTWKSCKCPMGFSYDT 2640
2641 RCFDSTVTESDIRVEESIYQCCLDAPARQAIRSLTERLYIGGLPLTNSKGQNGCYRRCA 2700
2641 RCFDSTVTESDIRVEESIYQCCLDAPARQAIRSLTERLYIGGLPLTNSKGQNGCYRRCA 2700
2701 SGVLTTCGNTLTCYLKATAACRAAKQDCTMLVNGDDLTVICSSAGTQEDAAALRAFTE 2760
2701 SGVLTTCGNTLTCYLKATAACRAAKQDCTMLVNGDDLTVICSSAGTQEDAAALRAFTE 2760
2761 AMTRYSAPPGPPQPEYDLELITCSSNVSVAHASGRVYVYLTDRDPTTLARAAWETAR 2820
2761 AMTRYSAPPGPPQPEYDLELITCSSNVSVAHASGRVYVYLTDRDPTTLARAAWETAR 2820
2821 HTPINSWLGNIIMVAPTLWARMILMTHEFFSILLAEQLEKALDQIYGACYSIEPLDLPQ 2880
2821 HTPINSWLGNIIMVAPTLWARMILMTHEFFSILLAEQLEKALDQIYGACYSIEPLDLPQ 2880
2881 IIERLHGLSAFTLHVSYPGEINRVASCLRKLGVPPLRTWRRHARSYVRKLLSQGGAATC 2940

2881 IIERLHGLSAFTLHVSYPGEINRVASCLRKLGVPPLRTWRRHARSYVRKLLSQGGAATC 2940
2941 GRYLFNFWAVRTKLTPPIPAASQDLGSGWFWAGYSGGDIYHLSRARPWRPFPLCLLLLSV 3000
2941 GRYLFNFWAVRTKLTPPIPAASQDLGSGWFWAGYSGGDIYHLSRARPWRPFPLCLLLLSV 3000
3001 GVGIVYLLPNR 3010
3001 GVGIVYLLPNR 3010
RESULT 2
O92970_9HEPC
ID O92970_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O92970;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL Virology 244:161-172(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL: AF054248; AAC15723.1; -; Genomic_RNA.
DR PIR: A61196; A61196.
DR PIR: PQ0246; PQ0246.
DR PIR: PS0329; PS0329.
DR HSP: Q02828; INB4.
DR SW: Q92970; 1029-1657, 2008-2170, 2420-2949.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005324; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transmembrane; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRp.
DR InterPro: IPR002518; Pept_U39 HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_Ps.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV_capsid; 1.

DR	pfam; PF01542; HCV core; 1.	
DR	pfam; PF01539; HCV env; 1.	
DR	pfam; PF01560; HCV NS1; 1.	
DR	pfam; PF01538; HCV NS2; 1.	
DR	pfam; PF02907; HCV NS3; 1.	
DR	pfam; PF01006; HCV NS4a; 1.	
DR	pfam; PF01001; HCV NS4b; 1.	
DR	pfam; PF01506; HCV NS5a; 1.	
DR	pfam; PF00998; RRP_3; 1.	
DR	SMART; SM00487; DEXDC; 1.	
KW	Polyprotein.	
SQ	SEQUENCE 3010 AA; 326845 MW; 45FB399AD6141A88 CRC64;	
Query Match		
Best Local Similarity 99.6%; Score 15949; DB 2; Length 3010;		
Matches 2998; Conservative 7; Mismatches 5; Indels 0; Gaps 0;		
Qy	1 MSTNPKPQKTKRNTNRRPQDKVFPGGQIVGGVILLPRGPRLGVRATRKASERSQPRG 60	
Db	1 MSTNPKPQKTKRNTNRRPQDKVFPGGQIVGGVILLPRGPRLGVRATRKASERSQPRG 60	
Qy	61 RROPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPGRSRPSWGPTDPRRSRLNG 120	
Db	61 RROPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPGRSRPSWGPTDPRRSRLNG 120	
Qy	121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180	
Db	121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180	
Qy	181 LLSCLTIPASAVEVRNVSIGIYHVTNDCNSSIIVYEAADVIMHTPGCPVCQEGNSSRCWV 240	
Db	181 LLSCLTIPASAVEVRNVSIGIYHVTNDCNSSIIVYEAADVIMHTPGCPVCQEGNSSRCWV 240	
Qy	241 ALTPTLAARNASVPTTIIRRHVDLLVGTAAAPCSAMVYVGLDGSIFLVSQLFTFSPRRHET 300	
Db	241 ALTPTLAARNASVPTTIIRRHVDLLVGTAAAPCSAMVYVGLDGSIFLVSQLFTFSPRRHET 300	
Qy	301 VQDCNCSIPYGHVSGHRMADMMNNSPTTALVVSOLLRIPOAVDMDVAGAHGVLGLA 360	
Db	301 VQDYNCSIPYGHVSGHRMADMMNNSPTTALVVSOLLRIPOAVDMDVAGAHGVLGLA 360	
Qy	361 YYSWGNWAKVLIALLFAGVDGETHTCGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW 420	
Db	361 YYSWGNWAKVLIALLFAGVDGETHTCGRVAGHTTSGFTSLFSSGASQKIQLVNTNGSW 420	
Qy	421 HINRTALNCNDSIQTGFFAALFYAHKFNSSGCCPERMACRPIIDWFAGQGWGPIITYTKPNSS 480	
Db	421 HINRTALNCNDSIQTGFFAALFYAHKFNSSGCCPERMACRPIIDWFAGQGWGPIITYTKPNSS 480	
Qy	481 DORPYCWHYAPRCGVVPASQVCGPYCYCTPSPVNVGTTDRSGVPTYSGENETDVMILN 540	
Db	481 DORPYCWHYAPRCGVVPASQVCGPYCYCTPSPVNVGTTDRSGVPTYSGENETDVMILN 540	
Qy	541 NTRPPQGNFGCTWMNSTGFTKCGGPPNCNIGGVGNRTLICPTDCPRKPEATYTKCGSG 600	
Db	541 NTRPPQGNFGCTWMNSTGFTKCGGPPNCNIGGVGNRTLICPTDCPRKPEATYTKCGSG 600	
Qy	601 PWLTPRCLVDYPYRLWHYPCTLNFSIFPKVMYVGGVEHRLNAAACNTRGERCNLEDRDRS 660	
Db	601 PWLTPRCLVDYPYRLWHYPCTLNFSIFPKVMYVGGVEHRLNAAACNTRGERCNLEDRDRS 660	
Qy	661 ELSPLLLSTTEWQILPCAFTTLTALSTGLIHLHQNIVDVQYLYGVGSFAFSAIKWEYIL 720	
Db	661 ELSPLLLSTTEWQILPCAFTTLTALSTGLIHLHQNIVDVQYLYGVGSFAFSAIKWEYIL 720	
Qy	721 LFLLLADARVCACLMWMLIIAQEAALENLVINAASVAGAHGILSFLVFFCAAWYIKG 780	
Db	721 LFLLLADARVCACLMWMLIIAQEAALENLVINAASVAGAHGILSFLVFFCAAWYIKG 780	
Qy	781 RLAPGAAYAFYGVWPLLLLLLALPPRAYALDREMAASCAGVALVGLVFLTSLSPYKVFLLT 840	
Db	781 RLAPGAAYAFYGVWPLLLLLLALPPRAYALDREMAASCAGVALVGLVFLTSLSPYKVFLLT 840	

Qy	841 RLIIWMLQYFITRAEAHMQVWVPPVLNVRGGRDAIILLTCAVHPELIFDITKLLALLAILGLPLM 900	
Db	841 RLIIWMLQYFITRAEAHMQVWVPPVLNVRGGRDAIILLTCAVHPELIFDITKLLALLAILGLPLM 900	
Qy	901 VLQAGITRVYFVRAOGLIRACMLVRKVAGHYVQMVKLGALTGTVVYVHNLTPLRDWA 960	
Db	901 VLQAGITRVYFVRAOGLIRACMLVRKVAGHYIQMAFMKLGALTGTVVYVHNLTPLRDWA 960	
Qy	961 HAGLRDLAVAVEPVVFSAMETKVLITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020	
Db	961 HAGLRDLAVAVEPVVFSAMETKVLITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020	
Qy	1021 QGWRLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVVSSTATQSFATCINGCWT 1080	
Db	1021 QGWRLAPITAYSQOTRGVLGCVITSLTGRDKNOVEGEVQVVSSTATQSFATCINGCWT 1080	
Qy	1081 VYHAGSKTLAGPKGPIITQMTYTNVDLILVHQAPPGARSMTPCSCGSSDLVLTVRHADVI 1140	
Db	1081 VYHAGSKTLAGPKGPIITQMTYTNVDLILVHQAPPGARSMTPCSCGSSDLVLTVRHADVI 1140	
Qy	1141 PVRRGDSRGSLLSPRPVSYLKGSGGELLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200	
Db	1141 PVRRGDSRGSLLSPRPVSYLKGSGGELLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200	
Qy	1201 METTMRSPVFTDNSTPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQGYKVLVLPNSVAA 1260	
Db	1201 METTMRSPVFTDNSTPPAVPQTFQVAHLHAPTSGSKSTKVPAAVAAQGYKVLVLPNSVAA 1260	
Qy	1261 TLGFGAYMSKAHGIDPNIRITGVRTITTCGSIITYSTYSGKFLADGGCGGGAYDIIICDECHS 1320	
Db	1261 TLGFGAYMSKAHGIDPNIRITGVRTITTCGSIITYSTYSGKFLADGGCGGGAYDIIICDECHS 1320	
Qy	1321 TDSITLIGIGTVLDOAETAGARLVVLTATPPGSTVTVPHNIEELGLSNGGSIPIYGVKAI 1380	
Db	1321 TDSITLIGIGTVLDOAETAGARLVVLTATPPGSTVTVPHNIEELGLSNGGSIPIYGVKAI 1380	
Qy	1381 PIEALIKGRHLIFCHSKKKKDELAALGTGLGNAVAYVYRGLDVSVIPPTGDIVVATDAL 1440	
Db	1381 PIEALIKGRHLIFCHSKKKKDELAALGTGLGNAVAYVYRGLDVSVIPPTGDIVVATDAL 1440	
Qy	1441 MTGFTGDFSDVIDNCNCTVQTQVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGSGIYR 1500	
Db	1441 MTGFTGDFSDVIDNCNCTVQTQVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGSGIYR 1500	
Qy	1501 FVTPGERSGMPDSSVLCEDVAGCAWYELTTPAETSRLRAYLNTPLGLPVCODHLEFWES 1560	
Db	1501 FVTPGERSGMPDSSVLCEDVAGCAWYELTTPAETSRLRAYLNTPLGLPVCODHLEFWES 1560	
Qy	1561 VFTGLTHIDAHFLSOTKAGDNFPYLVAQATVCARAQAPPPSWQMWKCLIRLKPTLHG 1620	
Db	1561 VFTGLTHIDAHFLSOTKAGDNFPYLVAQATVCARAQAPPPSWQMWKCLIRLKPTLHG 1620	
Qy	1621 PTPLLYRLGAVQNEVILTHPIITKYIMACMSADLEVVVSTWVLVGGVLAALAAAYCLTTGVS 1680	
Db	1621 PTPLLYRLGAVQNEVILTHPIITKYIMACMSADLEVVVSTWVLVGGVLAALAAAYCLTTGVS 1680	
Qy	1681 VIVGRIILISGKPAVVDREVLYQEBDEMEECASOLPYTEQGMOLAEQFKKALGLLOAT 1740	
Db	1681 VIVGRIILISGKPAVVDREVLYQEBDEMEECASOLPYTEQGMOLAEQFKKALGLLOAT 1740	
Qy	1741 KQAEAAAAPVESKRWALETFWAKHWNFIISGQYLAGLSTLPCNPAIASLMAFTASITSP 1800	
Db	1741 KQAEAAAAPVESKRWALETFWAKHWNFIISGQYLAGLSTLPCNPAIASLMAFTASITSP 1800	
Qy	1801 LTTQNTLLFNILGGWVAQAQAPPSAASAFVAGIAGAAVGSIGLGKVLVDIILAGYGAGVA 1860	
Db	1801 LTTQNTLLFNILGGWVAQAQAPPSAASAFVAGIAGAAVGSIGLGKVLVDIILAGYGAGVA 1860	
Qy	1861 GALVAFKVMGSEVPSTEDLVNLLPALISPGALVGVVCAAILRRHVGPGEVQVWNNRLI 1920	
Db	1861 GALVAFKVMGSEVPSTEDLVNLLPALISPGALVGVVCAAILRRHVGPGEVQVWNNRLI 1920	
Qy	1921 AFASRGNHVSPTHVYPESDAAAARVTQIILSSLTITQLLKLHQLHWINEDCSTPCSGSWLRDV 1980	

Db 1921 AFASRGNHVSPTHYVPESDAARVQTQLSSLTITQLLKLHOWINEDCSTPCSGSWLRDV 1980
1981 WDWICTVLTDFTKTLQSKLLPRLPGVPFLSCORGKYGVRGDCIMOTTCCPCAQIAGHYK 2040
1981 WDWICTVLTDFTKTLQSKLLPRLPGVPFLSCORGKYGVRGDCIMOTTCCPCAQIAGHYK 2040
2041 NGSMRIVGRPTCSNTWHGTFPPINAYTTGCTSPSPAENYSGRALWRVAABEYVEVTRVGDH 2100
2041 NGSMRIVGRPTCSNTWHGTFPPINAYTTGCTSPSPAENYSGRALWRVAABEYVEVTRVGDH 2100
2101 YVTGTTDNVCKPCQVPAPEFFTEVDGVRHLRVAPACKPLLRDVTFOVLNLYLVSOL 2160
2101 YVTGTTDNVCKPCQVPAPEFFTEVDGVRHLRVAPACKPLLRDVTFOVLNLYLVSOL 2160
2161 PCEPEPDVTLTSMLTDPSSHITAETAKRLRLARGSPPSLASSASQLSAPLSKATCTTHD 2220
2161 PCEPEPDVTLTSMLTDPSSHITAETAKRLRLARGSPPSLASSASQLSAPLSKATCTTHD 2220
2221 SPDADLIEANLLWRQBMGNITRVESKVVILDSFEPLHAEGDEREISVAABILKRSK 2280
2221 SPDADLIEANLLWRQBMGNITRVESKVVILDSFEPLHAEGDEREISVAABILKRSK 2280
2281 FPSALPIWARPDPYNPFLLESWKDPDYVPVHVGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
2281 FPSALPIWARPDPYNPFLLESWKDPDYVPVHVGCPLPPTKAPPIPPRRKRTVVLTESNV 2340
2341 SSALAEATKTFGSSGSSAVDSGTATPLASDDGKGDVDSYSMSMPLEGEPPDPL 2400
2341 SSALAEATKTFGSSGSSAVDSGTATPLASDDGKGDVDSYSMSMPLEGEPPDPL 2400
2401 SDGSWTSVSEASEDVVCCSMSYTWTCALITPCAABESKLPINPLNSLLRHHNMVYAT 2460
2401 SDGSWTSVSEASEDVVCCSMSYTWTCALITPCAABESKLPINPLNSLLRHHNMVYAT 2460
2461 SRASLRQKVTDRLOVLDHVRDVLKEMKASTVKAKLIEEACKLTTPHSAKSKF 2520
2461 SRASLRQKVTDRLOVLDHVRDVLKEMKASTVKAKLIEEACKLTTPHSAKSKF 2520
2521 GYKAKVRNLSRAVNHRSVWEDLLEDTETPTDITIMAKSEVFCVQPEKGRKPARLIV 2580
2521 GYKAKVRNLSRAVNHRSVWEDLLEDTETPTDITIMAKSEVFCVQPEKGRKPARLIV 2580
2581 FPDGLVRCVKMALYDVVSTLPQAVMGSSYGFQSPKQRFVFLVNTWKSCKPMGFSDYT 2640
2581 FPDGLVRCVKMALYDVVSTLPQAVMGSSYGFQSPKQRFVFLVNTWKSCKPMGFSDYT 2640
2641 RCFDSTVTSIDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNCYRRCA 2700
2641 RCFDSTVTSIDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNCYRRCA 2700
2701 SGVLTSCGNTLTCYLKATACRAAKLODCTMLVNGDDLVIICBSAGTQSDAAALRAFTE 2760
2701 SGVLTSCGNTLTCYLKATACRAAKLODCTMLVNGDDLVIICBSAGTQSDAAALRAFTE 2760
2761 AMTRYSAPPDPPPOPEYDLELITSCSNVSVVAHDASGRVYVYLRDPTTPLARAWEATAR 2820
2761 AMTRYSAPPDPPPOPEYDLELITSCSNVSVVAHDASGRVYVYLRDPTTPLARAWEATAR 2820
2821 HTPINSWLGNIIMVAPTLNARMILMTWTFPSILLAQEQLEKALDCQIYGCACYSIEPLDLPQ 2880
2821 HTPINSWLGNIIMVAPTLNARMILMTWTFPSILLAQEQLEKALDCQIYGCACYSIEPLDLPQ 2880
2881 IIERLHGLSAPTLHSYSPGEINRVASCLRLKGVPLTRWRRARSVRAKLISQCGRAATC 2940
2881 IIERLHGLSAPTLHSYSPGEINRVASCLRLKGVPLTRWRRARSVRAKLISQCGRAATC 2940
2941 GRYLFWAVRTKLLTPTPAASQLDLSGFWVAGYSGGDIYHSLGRAPRWFPCLCLLLSV 3000
2941 GRYLFWAVRTKLLTPTPAASQLDLSGFWVAGYSGGDIYHSLGRAPRWFPCLCLLLSV 3000
3001 GVGIVLLPNR 3010
|||||

Db 3001 GVGIVLLPNR 3010
RESULT 3
O92971_9HEPC PRELIMINARY; PRT: 3010 AA.
AC O92971_9HEPC
DT 01-NOV-1998 (TREMELREL. 08, Created)
DT 01-NOV-1998 (TREMELREL. 08, Last sequence update)
DT 01-MAR-2004 (TREMELREL. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxid=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL Virology 244:161-172(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshina M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Seed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL: AF054249; AAC15724.1; -; Genomic_RNA.
DR PIR: A61196; A61196.
DR PIR: P00246; P00246.
DR PIR: P0804; P0804.
DR PIR: P0329; P0329.
DR HSSP: O02828; INB4.
DR SMR: O92971; 1029-1657, 2008-2170, 2420-2949.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR004109; peptidase_S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.

QY 1921 AFASRGNHVSPTHYVPESDAAAVTQILSSLTITQLLKHLOWINBDCSTPCSGSWLRDV 1980
DB 1921 AFASRGNHVSPTHYVPESDAAAVTQILSSLTITQLLKHLOWINBDCSTPCSGSWLRDV 1980
QY 1981 WDWICTVLDTFKTWLQSKLLPRLPGVPPFLSCQRYGKVGWRGDMOTTCPCGAQIAGHVK 2040
DB 1981 WDWICTVLDTFKTWLQSKLLPRLPGVPPFLSCQRYGKVGWRGDMOTTCPCGAQIAGHVK 2040
QY 2041 NGSMRIVGPRTCNTHWGTFPIINAYTTGCTPPAPNYSRALNRVAEEVEVTRVGDFH 2100
DB 2041 NGSMRIVGPRTCNTHWGTFPIINAYTTGCTPPAPNYSRALNRVAEEVEVTRVGDFH 2100
QY 2101 YVTGMTTNDVVKCQVPAPEFFTEVDGVRHLRYAPACKPRLLEDVTFQVGLNQLYVGSOL 2160
DB 2101 YVTGMTTNDVVKCQVPAPEFFTEVDGVRHLRYAPACKPRLLEDVTFQVGLNQLYVGSOL 2160
QY 2161 PCPEPDVTVLTSMLTDPSSHITAEAKRLARGSPPLSSASSASOLSAPLSKATCTTHHD 2220
DB 2161 PCPEPDVTVLTSMLTDPSSHITAEAKRLARGSPPLSSASSASOLSAPLSKATCTTHHD 2220
QY 2221 SPADLIEANLLRQEMGGNITRVESKNVILDSPEPLHAEGDEREISVAAEILKRSRK 2280
DB 2221 SPADLIEANLLRQEMGGNITRVESKNVILDSPEPLHAEGDEREISVAAEILKRSRK 2280
QY 2281 FPSALPIWARPDYNPPLLESWKDPDYVPPVHGCPPLPPTKAPPIPPRRKRTVVLTESNV 2340
DB 2281 FPSALPIWARPDYNPPLLESWKDPDYVPPVHGCPPLPPTKAPPIPPRRKRTVVLTESNV 2340
QY 2341 SSALAEIATKTFGSSGSADVSGTATALPDLSADGDKGSDVESYSNMPLEGPDPDL 2400
DB 2341 SSALAEIATKTFGSSGSADVSGTATALPDLSADGDKGSDVESYSNMPLEGPDPDL 2400
QY 2401 SDGSWSVTSBEASDVVCCSNYSYTWGALITPCAAEESKLPINPLSNLLRHHNMVYATT 2460
DB 2401 SDGSWSVTSBEASDVVCCSNYSYTWGALITPCAAEESKLPINPLSNLLRHHNMVYATT 2460
QY 2461 SRSASLRQKQVTFRLQVLDDHYRDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
DB 2461 SRSASLRQKQVTFRLQVLDDHYRDLVKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
QY 2521 GYGAKVRLNLSRAVNHRSWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
DB 2521 GYGAKVRLNLSRAVNHRSWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
QY 2581 FPDGLGVRCEKVALYDVVSTLPOAVMGSSYCFQYSPKQVVEFLVNTWKSCKPCMGFSYDT 2640
DB 2581 FPDGLGVRCEKVALYDVVSTLPOAVMGSSYCFQYSPKQVVEFLVNTWKSCKPCMGFSYDT 2640
QY 2641 RCFDSTVTESDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNCGYRRRA 2700
DB 2641 RCFDSTVTESDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNCGYRRRA 2700
QY 2701 SGVLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVCESAGTQEDAAALRAPTE 2760
DB 2701 SGVLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVCESAGTQEDAAALRAPTE 2760
QY 2761 AMTRYSAAPPDPPQPEYDLELITSCSSNVSVAHADSGKRVYLLTRDPTTTLARAWEYAR 2820
DB 2761 AMTRYSAAPPDPPQPEYDLELITSCSSNVSVAHADSGKRVYLLTRDPTTTLARAWEYAR 2820
QY 2821 HTPINSWLGNIIMVAPTILWARMILMTHFFSILLAQEKLKALDCQIYGACYSIEPLDLQP 2880
DB 2821 HTPINSWLGNIIMVAPTILWARMILMTHFFSILLAQEKLKALDCQIYGACYSIEPLDLQP 2880
QY 2881 IIERLHGLSAPFLHSYSPEINRVASCLRLKLGVPPLRTWRHRSVRKALLSQGGAATC 2940
DB 2881 IIERLHGLSAPFLHSYSPEINRVASCLRLKLGVPPLRTWRHRSVRKALLSQGGAATC 2940
QY 2941 GRYLFNWAVRTKLKLTPIPAASQLDLSGWVFAGYSGGDIYHSLSRARPRWPLCLLLLSV 3000
DB 2941 GRYLFNWAVRTKLKLTPIPAASQLDLSGWVFAGYSGGDIYHSLSRARPRWPLCLLLLSV 3000

QY 3001 GVGIVLLPNR 3010
DB 3001 GVGIVLLPNR 3010
RESULT 4
Q02828_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q02828;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RA MEDLINE=92391112; PubMed=1325713; Yoshizawa H., Iizuka H., Tanaka T.,
RX Okamoto H., Kojima M., Okada S., Yamamoto O., Sho Y. et al.;
RA Muchmore E.E., Peterson D.A., Ito Y., Mishiro S.;
RT "Genetic drift of hepatitis C virus during an 8.2-year infection in a
RT chimpanzee: variability and stability.";
RL Virology 190:894-899(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.;
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Sho Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; D10750; BAA01583.1; -; Genomic_RNA.
DR PIR; A61196; A61196
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR PDB; 1NB4; X-ray; A/B=2420-2989.
DR PDB; 1NB6; X-ray; A/B=2420-2989.
DR PDB; 1NB7; X-ray; A/B=2420-2989.
DR SNR; Q02828; 1029-1657, 2008-2170.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003688; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.

Db 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPALVGVVCAALRRHHVGPGEAVQWNRLLI 1920
Qy 1921 AFASRGHVSPTHYVPESDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSGSLRDV 1980
Db 1921 AFASRGHVSPTHYVPESDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSGSLRDV 1980
Qy 1981 WDWICTVLTDPKTLWQSLKLPRLPGVPFLSCQRYGKVGWGDGIMQTTCCGQAQIAGHVK 2040
Db 1981 WDWICTVLTDPKTLWQSLKLPRLPGVPFLSCQRYGKVGWGDGIMQTTCCGQAQIAGHVK 2040
Qy 2041 NGSMRIIVGPRTCNTWHTGTFPIINAYTTGCTPPAPNYSRALNRVAEEVVEVTRVGDHF 2100
Db 2041 NGSMRIIVGPRTCNTWHTGTFPIINAYTTGCTPPAPNYSRALNRVAEEVVEVTRVGDHF 2100
Qy 2101 YVTGMTTDNVKPCQVPAPEFFTEVDGVRHLHRYPACKPRLLEDVTFQVGLNQYLVGSOL 2160
Db 2101 YVTGMTTDNVKPCQVPAPEFFTEVDGVRHLHRYPACKPRLLEDVTFQVGLNQYLVGSOL 2160
Qy 2161 PCBPEDVTVLTSMLTDPSSHITABTAKRRLARGSPPSLASSASOLSAPSLKATCTTHHD 2220
Db 2161 PCBPEDVTVLTSMLTDPSSHITABTAKRRLARGSPPSLASSASOLSAPSLKATCTTHHD 2220
Qy 2221 SPADALTEANLLWFOEMGNITRVESENKVILDSPEPLHAEGDEREISVAAEILKRSRK 2280
Db 2221 SPADALTEANLLWFOEMGNITRVESENKVILDSPEPLHAEGDEREISVAAEILKRSRK 2280
Qy 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTSNV 2340
Db 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTSNV 2340
Qy 2341 SSALAEATKTFGSGSASVDSGTATLPDLASDDGKSGDSVESYSSMPLEGPDPDL 2400
Db 2341 SSALAEATKTFGSGSASVDSGTATLPDLASDDGKSGDSVESYSSMPLEGPDPDL 2400
Qy 2401 SDGSWSTVSEASEDVCSSNSYTWGALITPCAAEESKLPINPLSNLLRHHNMVATT 2460
Db 2401 SDGSWSTVSEASEDVCSSNSYTWGALITPCAAEESKLPINPLSNLLRHHNMVATT 2460
Qy 2461 SRSASLRQKVTFRLOLVLDHYRDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
Db 2461 SRSASLRQKVTFRLOLVLDHYRDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSF 2520
Qy 2521 GYGAKDVRLSSRAVNHRSWEDLLETPIPTTINAKSEVFCVPEKGGKRPARLIV 2580
Db 2521 GYGAKDVRLSSRAVNHRSWEDLLETPIPTTINAKSEVFCVPEKGGKRPARLIV 2580
Qy 2581 FPDGLGRVCEKVALYDVVSTLPOAVMGSSYGFQYSPKQVVEFLVNTWKSCKPWFESYDT 2640
Db 2581 FPDGLGRVCEKVALYDVVSTLPOAVMGSSYGFQYSPKQVVEFLVNTWKSCKPWFESYDT 2640
Qy 2641 RCFDSTVTESDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNCGYRRCRA 2700
Db 2641 RCFDSTVTESDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNCGYRRCRA 2700
Qy 2701 SGVLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRAFTE 2760
Db 2701 SGVLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALRAFTE 2760
Qy 2761 AMTRYSAFPQPPQPEYDLELITSCSNVSVVAHDASGRVYVLTDRDPTTLARAWEATAR 2820
Db 2761 AMTRYSAFPQPPQPEYDLELITSCSNVSVVAHDASGRVYVLTDRDPTTLARAWEATAR 2820
Qy 2821 HTPINSWLGNIIMYAPTILWARMILMTHFFSILLAOEKLDCQIYGACYSIEPLDLQP 2880
Db 2821 HTPINSWLGNIIMYAPTILWARMILMTHFFSILLAOEKLDCQIYGACYSIEPLDLQP 2880
Qy 2881 IIERLHGLSAFTLHYSYSGEINRVASCLRLKGLVPLTRWRHARSVRKLLSQGGRAATC 2940
Db 2881 IIERLHGLSAFTLHYSYSGEINRVASCLRLKGLVPLTRWRHARSVRKLLSQGGRAATC 2940
Qy 2941 GRYLFNWAVRTKLKLTPIPAASQLDLSGNFVAGYSGGDIYHSLSRARPRWPLCLLLLSV 3000
Db 2941 GRYLFNWAVRTKLKLTPIPAASQLDLSGNFVAGYSGGDIYHSLSRARPRWPLCLLLLSV 3000

Qy 3001 GVGIVLLPNR 3010
Db 3001 GVGIVLLPNR 3010

RESULT 5

O92972_9HEPC
ID O92972_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O92972;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo.";
RL Virology 244:161-172(1998).
RN [2]
NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Ohnima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL: AF054250; AAC15725.1; -; Genomic_RNA.
DR PIR: A61196; A61196.
DR PIR: PQ0246; PQ0246.
DR PIR: PS0329; PS0329.
DR HSPR: Q02828; INB4.
DR SMR: O92972; 1029-1657, 2008-2170, 2420-2949.
DR GO: 0019028; C:viral capsid; IEA.
DR GO: 0019031; C:viral envelope; IEA.
DR GO: 0005524; F:ATP binding; IEA.
DR GO: 0008026; F:ATP-dependent helicase activity; IEA.
DR GO: 0003723; F:RNA binding; IEA.
DR GO: 0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: 0008236; F:serine-type peptidase activity; IEA.
DR GO: 0005198; F:structural molecule activity; IEA.
DR GO: 0006508; F:proteolysis and peptidolysis; IEA.
DR GO: 0006350; P:transcription; IEA.
DR GO: 0019079; P:viral genome replication; IEA.
DR GO: 0019087; P:viral transformation; IEA.
DR InterPro: IPR001410; DEAD
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV_NSI.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR002518; Pept_U39 HCV NS2.
DR InterPro: IPR004109; peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_Psvir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV_NSI; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.

Db	1981	WDMICTVLTDFKTMWLOSKLLRPLRGVVFLLSCQGYKGWNRGDGIMQTTCPGGAQIAGHVK	2040	ID	Q02829_9HEPC PRELIMINARY;	PRT;	3010	AA.
Qy	2041	NGSMRIVGPRTCNTNTHGTFPINAYTTGCTPSPAPNYSRALNRVAEEYVEVTRVGDHF	2100	AC	Q02829;			
Db	2041	NGSMRIVGPRTCNTNTHGTFPINAYTTGCTPSPAPNYSRALNRVAEEYVEVTRVGDHF	2100	DT	01-NOV-1996 (TREMBLrel. 01, Created)			
Qy	2101	YVTGMTDNVKKCOVPAPPEFFTEVDGRLRHRYAPACKPLLRDVTFOVGLNOYLVSQ	2160	DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
Db	2101	YVTGMTDNVKKCOVPAPPEFFTEVDGRLRHRYAPACKPLLRDVTFOVGLNOYLVSQ	2160	DE	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
Qy	2161	PCPEPDVTVLTSMLTDPSSHITAETAKRLRLARGSPPSLASSASQSLSPSKATCTTHHD	2220	OS	Hepatitis C virus.			
Db	2161	PCPEPDVTVLTSMLTDPSSHITAETAKRLRLARGSPPSLASSASQSLSPSKATCTTHHD	2220	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;			
Qy	2221	SPDADLIEANLLRQEMGNITRVESENKVVILDSFEPLHAEGDEREISVAEILKRSK	2280	OC	Hepacivirus.			
Db	2221	SPDADLIEANLLRQEMGNITRVESENKVVILDSFEPLHAEGDEREISVAEILKRSK	2280	OX	NCBI_TaxID=11103;			
Qy	2281	FPSALPIWARPDYNPPLLESWKDDYVPVYVHGCPPLPTKAPPIPPRRKRTVVLTESNV	2340	RN	[1]			
Db	2281	FPSALPIWARPDYNPPLLESWKDDYVPVYVHGCPPLPTKAPPIPPRRKRTVVLTESNV	2340	RP	NUCLEOTIDE SEQUENCE.			
Qy	2341	SSALAEATATKFGSSGSAVDSTGATATPLDASDGDGKSDVESYSSMPLEGEPPDPL	2400	RC	STRAIN=HC-J4;			
Db	2341	SSALAEATATKFGSSGSAVDSTGATATPLDASDGDGKSDVESYSSMPLEGEPPDPL	2400	RX	MEDLINE=92391112; PubMed=13257113;			
Qy	2401	SDGSWSTVSEASEDDVCCSMSTYWTGALITPCAAEESKLPINPLNSLLRHNMVYATT	2460	RA	Okamoto H., Kojima M., Okada S., Yoshizawa H., Iizuka H., Tanaka T.,			
Db	2401	SDGSWSTVSEASEDDVCCSMSTYWTGALITPCAAEESKLPINPLNSLLRHNMVYATT	2460	RA	Muchmore E.E., Peterson D.A., Ito Y., Mishihiro S.;			
Qy	2461	SRASLRQKKVTFDRLOVLDHRYDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSKF	2520	RA	"Genetic drift of hepatitis C virus during an 8.2-year infection in a			
Db	2461	SRASLRQKKVTFDRLOVLDHRYDVLKEMKAKASTVKAKLLSTEEACKLTPPHSAKSKF	2520	RT	chimpanzee: variability and stability.";			
Qy	2521	GYGAKVRNLSRAVNHRSWEDLLEDTEPDIITIMAKSEVFCVQPEKGRKPARLIV	2580	RL	single patient show sequence heterogeneity.";			
Db	2521	GYGAKVRNLSRAVNHRSWEDLLEDTEPDIITIMAKSEVFCVQPEKGRKPARLIV	2580	RL	J. Gen. Virol. 72:2805-2809(1991).			
Qy	2581	FPDLGVRCEKMALYDVVSTLPQAVMGSSYGFQSPKQRFVEFLVNTWKSCKPMGFSDYT	2640	DR	EMBL; D13558; BAA02756.1; -; Genomic_RNA.			
Db	2581	FPDLGVRCEKMALYDVVSTLPQAVMGSSYGFQSPKQRFVEFLVNTWKSCKPMGFSDYT	2640	DR	PIR; A61196; A61196.			
Qy	2641	RCFDSVTESDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGONCYRRCA	2700	DR	PIR; PS0246; PS0246.			
Db	2641	RCFDSVTESDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLTNSKGONCYRRCA	2700	DR	HSSP; Q8YYS1; LCW.			
Qy	2701	SGVLTSCGNLTLCYLKATAACRAKLODCTMLVNGDDLVIICESAGTOEDAAALRAPTE	2760	DR	SMR; Q02829; 1029-1657, 2008-2170, 2420-2949.			
Db	2701	SGVLTSCGNLTLCYLKATAACRAKLODCTMLVNGDDLVIICESAGTOEDAAALRAPTE	2760	DR	GO; GO:0019028; C:Viral capsid; IEA.			
Qy	2761	AMTRYSAPPDPPOPEYDLELITSCSSNVSAHDASGKRVYLLTRDPTTPLARAAMETAR	2820	DR	GO; GO:0019031; C:Viral envelope; IEA.			
Db	2761	AMTRYSAPPDPPOPEYDLELITSCSSNVSAHDASGKRVYLLTRDPTTPLARAAMETAR	2820	DR	GO; GO:0005524; F:ATP binding; IEA.			
Qy	2821	HTPINSWLGNIIMVAPTLWARMLMTWFFSILLAQEKLDCQIYGACYSIPLDLQP	2880	DR	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.			
Db	2821	HTPINSWLGNIIMVAPTLWARMLMTWFFSILLAQEKLDCQIYGACYSIPLDLQP	2880	DR	GO; GO:0003723; F:RNA binding; IEA.			
Qy	2881	IIERLHGLSAPTLHSYSGEINRVASCLKLVGPPLTRWRHARSVRKLLSQGRAATC	2940	DR	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.			
Db	2881	IIERLHGLSAPTLHSYSGEINRVASCLKLVGPPLTRWRHARSVRKLLSQGRAATC	2940	DR	GO; GO:0008236; F:serine-type peptidase activity; IEA.			
Qy	2941	GRYLFNWAIVTKLKLTPIPAASQDLGSGFWAGYSGGDIYHSLSRARPRWFPCLLLLSV	3000	DR	GO; GO:0005198; F:structural molecule activity; IEA.			
Db	2941	GRYLFNWAIVTKLKLTPIPAASQDLGSGFWAGYSGGDIYHSLSRARPRWFPCLLLLSV	3000	DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
Qy	3001	GVGIYLLPNR 3010		DR	GO; GO:0006350; P:transcription; IEA.			
Db	3001	GVGIYLLPNR 3010		DR	GO; GO:0019079; P:Viral genome replication; IEA.			
				DR	GO; GO:0019087; P:Viral transformation; IEA.			
				DR	InterPro; IPR001410; DEAD.			
				DR	InterPro; IPR011545; DEAD/DEAH N.			
				DR	InterPro; IPR002522; HCV capsid.			
				DR	InterPro; IPR002521; HCV_Core.			
				DR	InterPro; IPR002519; HCV Env.			
				DR	InterPro; IPR002531; HCV NS1.			
				DR	InterPro; IPR000745; HCV NS4a.			
				DR	InterPro; IPR001490; HCV NS4b.			
				DR	InterPro; IPR002868; HCV NS5a.			
				DR	InterPro; IPR002166; HCV RdRp.			
				DR	InterPro; IPR002518; Pept_U39_HCV NS2.			
				DR	InterPro; IPR004109; Peptidase_S29.			
				DR	InterPro; IPR007095; RNA_pol_DS_PS.			
				DR	InterPro; IPR007094; RNA_pol_psvir.			
				DR	Pfam; PF01543; HCV capsid; 1.			
				DR	Pfam; PF01542; HCV_Core; 1.			
				DR	Pfam; PF01539; HCV Env; 1.			
				DR	Pfam; PF01560; HCV NS1; 1.			
				DR	Pfam; PF01538; HCV NS2; 1.			
				DR	Pfam; PF02907; HCV NS3; 1.			
				DR	Pfam; PF01006; HCV NS4a; 1.			
				DR	Pfam; PF01001; HCV NS4b; 1.			
				DR	Pfam; PF01506; HCV NS5a; 1.			
				DR	Pfam; PF00998; RdRp_3; 1.			
				DR	SMART; SM00487; DEXDc; 1.			
				KW	Polyprotein.	1	191	C protein.
				FT	CHAIN	192	383	E protein.

FT	CHAIN	384	729	NS1/E2 protein.
FT	CHAIN	730	1006	NS2 protein.
FT	CHAIN	1007	1615	NS3 protein.
FT	CHAIN	1616	2013	NS4 protein.
FT	CHAIN	2014	3010	NS5 protein.
SQ	SEQUENCE	3010 AA;	326962 MW;	93D465526F3EADP8 CRC64;
Query Match				
Best Local Similarity 98.3%; Pred. No. 0;				
Matches 2959; Conservative 19; Mismatches 32; Indels 0; Gaps 0;				
QY	1	MSTNPKPQKTKRNTNR	PDQVKFPGGQIVGGVYLLPRRGPRLGVRATRKASERSQPRG	60
DB	1	MSTNPKPQKTKRNTNR	PQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG	60
QY	61	RRQPTPKARRPRGR	AWAQPGYPWPLYGNEGLGWAAGWLLSPGRSRPSWGPDPDRRSRNLG	120
DB	61	WRQP1PKARRPRGR	AWAQPGYPWPLYGNEGLGWAAGWLLSPGRSRPSWGPDPDRRSRNLG	120
QY	121	KVIDTLTCGFADLMGYI	PLVGAPLGGARALAHGVRVLEGGVYATGNLPGCSFISIFLLA	180
DB	121	KVIDTLTCGFADLMGYI	PLVGAPLGGARALAHGVRVLEGGVYATGNLPGCSFISIFLLA	180
QY	181	LLSCLTIPASAEYVR	NSGIYHVYVNDCSNSSIYVEAADVIMHTPGCPVCVOEGNSSRCWV	240
DB	181	LLSCLTIPASAEYVR	NSGIYHVYVNDCSNSSIYVEAADVIMHTPGCPVCVREDNSSRCWV	240
QY	241	ALPTPLAARNASVPT	TIRRHVDLVGTAAFGCSAMVVGDLCSIFLVSQLFTFSPRRHET	300
DB	241	ALPTPLAARNASVPT	TIRRHVDLVGTAAFGCSAMVVGDLCSIFLVSQLFTFSPRRHET	300
QY	301	VQDCNSIYPGHVSGH	RMADMMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLGLA	360
DB	301	VQDCNSIYPGHVSGH	RMADMMNWSPTTALVVSQLLRIPOAVVDMVAGAHGVLGLA	360
QY	361	YYSWGNWAKVILVALL	FAGVDGTHHTGRVAGHTTSGFSLFSSGASOKIOLVNTNGSW	420
DB	361	YYSWGNWAKVILVALL	FAGVDGTHHTGRVAGHTTSGFSLFSSGASOKIOLVNTNGSW	420
QY	421	HINRTALNCNDSLOT	GFALFYAHKFNSSGCPERMASCRPIDWAQAGWGPITYTKPNSS	480
DB	421	HINRTALNCNDSLOT	GFALFYAHKFNSSGCPERMASCRPIDWAQAGWGPITYTEPDSP	480
QY	481	DQPYCWHYAPRCGV	VPASOVCGPYCPTPSPVVVGTITDRSGVPTYSWGENETDVMLLN	540
DB	481	DQPYCWHYAPRCGV	VPASOVCGPYCPTPSPVVVGTITDRSGVPTYSWGENETDVMLLN	540
QY	541	NTRPPQGNWFGCTW	MNSTGFTKTCGGPPCNIGGVGNRTLICPTDCPRKHPEATYTKCGSG	600
DB	541	NTRPPQGNWFGCTW	MNSTGFTKTCGGPPCNIGGVGNRTLICPTDCPRKHPEATYTKCGSG	600
QY	601	PWLTPRCLVDYVRL	MWHPCTNLNFSIFKVRMYVGGVEHRLNAAACNWRGERCNLEDRDRS	660
DB	601	PWLTPRCLVDYVRL	MWHPCTNLNFSIFKVRMYVGGVEHRLNAAACNWRGERCNLEDRDRS	660
QY	661	ELSPILLSTTEWQIIL	PCATFTLIPALSTGLIHLHONTVDVQYLYGVCSAFVSAIKWEYIL	720
DB	661	ELSPILLSTTEWQIIL	PCATFTLIPALSTGLIHLHONTVDVQYLYGVCSAFVSAIKWEYIL	720
QY	721	LLFLLADARVCACL	WMWLLIIAQEAALLENVLNNAASVAGAHGILSFLVFFCAAWYIKG	780
DB	721	LLFLLADARVCACL	WMWLLIIAQEAALLENVLNNAASVAGAHGILSFLVFFCAAWYIKG	780
QY	781	RLAPGAAYFYGYW	PILLLLALLPPRAYALDREMAASCGGAVLVGLVFTLTSPPYKVFLL	840
DB	781	RLAPGAAYFYGYW	PILLLLALLPPRAYALDREMAASCGGAVLVGLVFTLTSPPYKVFLL	840
QY	841	RLIWWLOYFITRAE	AHQWVPPPLNVRRGGDAIILLTCAVHPBELIFDITKLLAILGLPLM	900
DB	841	RLIWWLOYFITRAE	AHQWVPPPLNVRRGGDAIILLTCAVHPBELIFDITKLLAILGLPLM	900
QY	901	VLQAGITRPVYFV	RAOGLIRACMLVRKVAGGHVQVFMKLGALTGTGYVYNHLTPLRDWA	960

DB	901	VLQAGITRPVYFV	RAOGLIRACMLVRKVAGGHVQVFMKLGALTGTGYVYNHLTPLRDWA	960
QY	961	HAGLRDLAVAVEP	VVFSAMETKVITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG	1020
DB	961	HAGLRDLAVAVEP	VVFSAMETKVITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG	1020
QY	1021	QGWELLAPITAYS	QOQTRGVLCIITSITGRDKNKGVEQVVSTATQSFATCINGVCWT	1080
DB	1021	QGWELLAPITAYS	QOQTRGVLCIITSITGRDKNKGVEQVVSTATQSFATCINGVCWT	1080
QY	1081	VYHAGSKTLAGPK	GPITQMYTNVVDLIVGWQAPPGARSMTPCSCGSDLLVLVTRHADVI	1140
DB	1081	VYHAGSKTLAGPK	GPITQMYTNVVDLIVGWQAPPGARSMTPCSCGSDLLVLVTRHADVI	1140
QY	1141	PVRRGDSRGSLLS	PRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES	1200
DB	1141	PVRRGDSRGSLLS	PRPVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES	1200
QY	1201	METTMRSFVFTDN	STPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQYKVLVLPNSVAA	1260
DB	1201	METTMRSFVFTDN	STPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQYKVLVLPNSVAA	1260
QY	1261	TLGFGAYMSKAH	GIDPNIRTVRTITGGSIITYSTYKFLADGCGSGGAYDIIICDECHS	1320
DB	1261	TLGFGAYMSKAH	GIDPNIRTVRTITGGSIITYSTYKFLADGCGSGGAYDIIICDECHS	1320
QY	1321	TDSITLIGITVLD	QAEATAGARLVVLTATPPGSVTVPHPNIEIGISNNGEIPFYGKAI	1380
DB	1321	TDSITLIGITVLD	QAEATAGARLVVLTATPPGSVTVPHPNIEIGISNNGEIPFYGKAI	1380
QY	1381	PIEAIKGRHILFCH	SKKCKDELAALKTLGLNLNAVAYRGDLVSVIPIPGDVAVVATDAL	1440
DB	1381	PIEAIKGRHILFCH	SKKCKDELAALKTLGLNLNAVAYRGDLVSVIPIPGDVAVVATDAL	1440
QY	1441	MTGTGPDSDVID	CNTCVTQTFVDFSLDPTFTIETTTVPQDAVSQRGRGSGIYR	1500
DB	1441	MTGTGPDSDVID	CNTCVTQTFVDFSLDPTFTIETTTVPQDAVSQRGRGSGIYR	1500
QY	1501	FVTGERPSGMFDS	SVLCECYDAGCAWVELTTPAETSURLRAYLNTPGLPVCODHLEFPWES	1560
DB	1501	FVTGERPSGMFDS	SVLCECYDAGCAWVELTTPAETSURLRAYLNTPGLPVCODHLEFPWES	1560
QY	1561	VFTGLTHIDAHFL	SQTKAGDNFPVLVAYQATVCARAQPPSDWDMWKCLIRLKPTLHG	1620
DB	1561	VFTGLTHIDAHFL	SQTKAGDNFPVLVAYQATVCARAQPPSDWDMWKCLIRLKPTLHG	1620
QY	1621	PTPLLRYLGAVQ	NEVILTHPIITKYIMACMSADLEVTSTWVLVGGVLAALAAAYCLTTGVS	1680
DB	1621	PTPLLRYLGAVQ	NEVILTHPIITKYIMACMSADLEVTSTWVLVGGVLAALAAAYCLTTGVS	1680
QY	1681	VIVGRIILSGKPA	VVPDBREVLYQBFDEMEBCASQIPYIEQGMQLAEOFKQKALGLLOTAT	1740
DB	1681	VIVGRIILSGKPA	VVPDBREVLYQBFDEMEBCASQIPYIEQGMQLAEOFKQKALGLLOTAT	1740
QY	1741	KQAEAAAAPVSES	KWRALETFWAKMWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSP	1800
DB	1741	KQAEAAAAPVSES	KWRALETFWAKMWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSP	1800
QY	1801	LTQNTLLFNILG	GVAAQLAAPPAAAFVAGIAGAAVSGIGIKVLVIDILAGYGAGVA	1860
DB	1801	LTQNTLLFNILG	GVAAQLAAPPAAAFVAGIAGAAVSGIGIKVLVIDILAGYGAGVA	1860
QY	1861	GALVAFKVMSE	VEPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGEAGVQMMNRLLI	1920
DB	1861	GALVAFKVMSE	VEPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGEAGVQMMNRLLI	1920
QY	1921	AFASRGHVSPTH	YVPESDAAAARVTQILSSLTITOLLKRLHWINEDCSTPCSGSWLRDV	1980
DB	1921	AFASRGHVSPTH	YVPESDAAAARVTQILSSLTITOLLKRLHWINEDCSTPCSGSWLRDV	1980
QY	1981	WDWICTVLTD	FKTWLQSKLLPRLFGVPFLSCQRYKGYMRGDMIMQTTCPCGAQIAGHVK	2040
DB	1981	WDWICTVLTD	FKTWLQSKLLPRLFGVPFLSCQRYKGYMRGDMIMQTTCPCGAQIAGHVK	2040

QY	2041	NGSMRIVGPRCTCNTWHGTPPINAYTTGCTPSPAPNYSRALNRVAABEEVETRVGDHF	2100
Db	2041		
QY	2101	YVTGMTDNDVKCPQCPAPAEFFTEVDGVRHLRYPACKPDLREDVTFQVGLNOLYVGSOL	2160
Db	2101		
QY	2161	PCPEPEPDVTLTSMITDPSHITAEAKRLRLARGSPPSLASSASQLSAPLSKATCTTHHD	2220
Db	2161		
QY	2221	SPDADLEAMLLRWQMGNNITRVESENKVILDSFEPLHAEGDERBISVAABILKRSRK	2280
Db	2221		
QY	2281	FPSPALPIWARPDPNPPLESWKDPDYPVPVHGCPLPPTKAPPTPPRRKRTVVLTESNV	2340
Db	2281		
QY	2341	SSALAEATATKFGSSGSADVSGTATALPDLASDGDGSDVESYSSMPLEGEPPGDPDL	2400
Db	2341		
QY	2401	SDGSWSTVSEASEDVVCCSMTYTWTCALITPCAAEESKLPINPLSNLLRHHNMVYATT	2460
Db	2401		
QY	2461	SRASLRKQKVTDFRLQVLDHYRDVLKEMKAKASTVKAKLSTEEACKLTPPHSAKSKF	2520
Db	2461		
QY	2521	GYGAKDVNLSRAVNHRSWEDLEDTEPIDTTIMAKSEVFCVQPEKGRKPARLIIV	2580
Db	2521		
QY	2581	FPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPKQRFVEFLVNTWKSCKCPMGFSYDT	2640
Db	2581		
QY	2641	RCFDSVTVESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCA	2700
Db	2641		
QY	2701	SGVLTTSCGNTLTCVLKATAACRAKQDCTMLVNGDDLVIICBSAGTQEDAAALRAFTE	2760
Db	2701		
QY	2761	AMTRYSAPPDPPOPEYDLELITSCSSNVSVAHDSAGKRVYLYLTRDPTTPIARAAMETAR	2820
Db	2761		
QY	2821	HTPINSWLGNIIMVAPALWARMILWTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPO	2880
Db	2821		
QY	2881	IIERLHGLSAPLTHSYSGEINRVASCLRLGVPLRTWRHRASVRKAKLSQGGRAATC	2940
Db	2881		
QY	2941	GRYLFNNAVRTKLTPTIPAAQSOLDLSCGFVAGVSGGDIYHSLSRARPRWFPLCILLLSV	3000
Db	2941		
QY	3001	GVGIYLLPNR 3010	
Db	3001	GVGIYLLPNR 3010	

RESULT 7
Q4PKP9_9HEPC
ID Q4PKP9_9HEPC PRELIMINARY; PRT; 3010 AA.

AC	Q4PKP9;		
DT	13-SEP-2005 (TReMBLrel. 31, Created)		
DT	13-SEP-2005 (TReMBLrel. 31, Last sequence update)		
DT	13-SEP-2005 (TReMBLrel. 31, Last annotation update)		
DE	Polyprotein.		
OS	Hepatitis C virus.		
OC	Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Hepacivirus.		
OX	NCBI_taxid=11103;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Yeh C.-T.;		
RT	"Emergence of mutation clusters in the HCV genome during sequential		
RT	viral passages in Sip-L expressing cells";		
RL	Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; DQ071885; AAY81920.1; -; mRNA.		
KW	Polyprotein.		
SQ	SEQUENCE 3010 AA; 327152 MW; 39DCAD60317840EC CRC64;		
	Query Match 96.3%; Score 15422; DB 2; Length 3010;		
	Best Local Similarity 95.9%; Pred. No. 0;		
	Matches 2887; Conservative 52; Mismatches 71; Indels 0; Gaps 0;		
QY	1	1	1
Db	1	1	1
QY	61	61	61
Db	61	61	61
QY	121	121	121
Db	121	121	121
QY	181	181	181
Db	181	181	181
QY	241	241	241
Db	241	241	241
QY	301	301	301
Db	301	301	301
QY	361	361	361
Db	361	361	361
QY	421	421	421
Db	421	421	421
QY	481	481	481
Db	481	481	481
QY	541	541	541
Db	541	541	541
QY	601	601	601
Db	601	601	601
QY	661	661	661
Db	661	661	661
QY	721	721	721
Db	721	721	721

Db 721 LFLLLADARVACALWMMMLLIAQAEAALENVLVNAASVAGAHGILSLFVFFCAAWYIKG 780
Qy 781 RLAPGAAVAFYGVWPLLLLLALPPRAVALDEMAASCGGAVLVGLVFLTLSPYKVELT 840
Db 781 RLAPGAAVAFYGVWPLLLLLALPPRAVALDEMAASCGGAVLVGLVFLTLSPYKVELA 840
Qy 841 RLIIWLQVFIITRAEAHQWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGLPLM 900
Db 841 RLIIWLQVFIITRAEHLQWVPLNVRGGRDAVILLTCAIHPELIFDITKLLAILGLPLM 900
Qy 901 VLOAGITRVPYFVRAQGLIRACWLRKVAGGHYQVQWFMKLGALTCVTVYVNHLPRLDWA 960
Db 901 VLOAGITRVPYFVRAQGLIRACWLRKVAGGHYQVQWFMKLAALTCVTVYVNHLPRLDWA 960
Qy 961 HAGLRDLAVAVPVPVFSAMETKVIITWGDATACAGDIIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HAGLRDLAVAVPVPVFSAMETKVIITWGDATACAGDIIILGLPVSARRGKEIHLGPADSLEG 1020
Qy 1021 QGWRLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVYSTATQSFATCINGVCWT 1080
Db 1021 QGWRLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVYSTATQSFATCINGVCWT 1080
Qy 1081 VYHGAGSKTLAPGKPIITQMYTNVDLIDLVGWQAPPGARSMTPCSCGSSDLXLVTRHADVI 1140
Db 1081 VYHGAGSKTLAPGKPIITQMYTNVDLIDLVGWQAPPGARSMTPCSCGSSDLXLVTRHADVI 1140
Qy 1141 PVRRGDSRGSLLSPRPVSYLKGSSGGLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRGDSRGSLLSPRPVSYLKGSSGGLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200
Qy 1201 METTMRSPFTDNSTPPAVPQTFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLNPSVAA 1260
Db 1201 METTMRSPFTDNSTPPAVPQTFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLNPSVAA 1260
Qy 1261 TLGFGAYMSKAHGDIPNITGVRTITTGSIITYSTYKFLADGGCGSGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGDIPNITGVRTITTGSIITYSTYKFLADGGCGSGAYDIIICDECHS 1320
Qy 1321 TOSTTILGTGLVDOAETAGARLVLATATPGSVTVPHPNTEETGLSNNGBIIPYGKAI 1380
Db 1321 TOSTTILGTGLVDOAETAGARLVLATATPGSVTVPHPNTEETGLSNNGBIIPYGKAI 1380
Qy 1381 PIEAIKGRGRLHIFCHSKKKCDBLAALKLGLGLNAVAYRGLDVSIVPIPGDVVVVATDAL 1440
Db 1381 PIETIKGRGRLHIFCHSKKKCDBLAALKLGLGLNAVAYRGLDVSIVPTSGDVIIVATDAL 1440
Qy 1441 MTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTTETTPQDAVRSQRGRGRTGRSGIYR 1500
Db 1441 MTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTTETTPQDAVRSQRGRGRTGRSGIYR 1500
Qy 1501 FVTPGRPSGMPDSSVLCBCYDAGCAWYELTPAETSVRLAYLNTPGLPVCODHLEFWES 1560
Db 1501 FVTPGRPSGMPDSSVLCBCYDAGCAWYELTPAETSVRLAYLNTPGLPVCODHLEFWES 1560
Qy 1561 VFTGLTHIDAHFLSQTQKAGDNFPLYVAYQATVCARAQAPPPSWDMWKCLIRLKPRTLHG 1620
Db 1561 VFTGLTHIDAHFLSQTQKAGDNFPLYVAYQATVCARAQAPPPSWDMWKCLIRLKPRTLHG 1620
Qy 1621 PTPLLYRLGAVQNEVLTHPTKYIMACMSADLEVVTSTWVLVGGVLAALAYCLTGTGSV 1680
Db 1621 PTPLLYRLGAVQNEVTHPTKYIMACMAADLEVVTSTWVLVGGVLAALAYCLTGTGSV 1680
Qy 1681 VIVGRILISGKPAVDPREVILVCEPDEMEECASOLPYIEQGMOLAEOPKQKALGLLOTAT 1740
Db 1681 VIVGRILISGKPAIIPDREVILVCEPDEMEECASHLPYIEQGMOLAEOPKQKALGLLOTAT 1740
Qy 1741 KQAEAAAPVVEKSWRALETFWAKHMNFISGIQYLAGLSTLPCNPAIASIMAFITASITSP 1800
Db 1741 KQAEAAAPVVEKSWRTLEAFWANDMNFISGIQYLAGLSTLPCNPAIASIMAFITASITSP 1800
Qy 1801 LTTQNTLLFNILGGWVAQAOLAPPSAASAFVAGIAGAAVGSITGLKVLVDIILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGGWVAQAOLAPPSAASAFVAGIAGAAVGSITGLKVLVDIILAGYGAGVA 1860

Qy 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQMMNRLI 1920
Db 1861 GALVAFKIMGSEMPSSDVLNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQMMNRLI 1920
Qy 1921 AFASRGNHVSPTHVVPESDAAARVTQILSSLTITQLLKRHLQWINEBCSTPCSSGSLRDV 1980
Db 1921 AFASRGNHVSPTHVVPESDAAARVTQILSSLTITQLLKRHLQWINEBCSTPCSSGSLRDV 1980
Qy 1981 WDMICTVLTFKWLQSKLLPRLPGVPFLSCQRYGKGVWRGDGIMQTTCPCGAOIAGHVK 2040
Db 1981 WDMICTVLTFKWLQSKLLPRLPGVPFLSCQRYGKGVWRGDGIMQTTCPCGAOIAGHVK 2040
Qy 2041 NGSNRIVGPRTCSNTHGTFPINAITYTGPCTPSPAPNYSRALMRVAAEYVEVTRVGDGFH 2100
Db 2041 NGSNRIVGPRTCSNTHGTFPINAITYTGPCTPSPAPNYSRALMRVAAEYVEVTRVGDGFH 2100
Qy 2101 YVTGMTDNVKKPCQVPAPPEFFTEVDGVRHRYAPACKPLLRREDVTQVGLNQYLVGSOL 2160
Db 2101 YVTGMTDNVKKPCQVPAPPEFFTEVDGVRHRYAPACKPLLRREDVTQVGLNQYLVGSOL 2160
Qy 2161 PCEPEPDVTVLTSMLTDPSHITAEAKRRLARGSPPSLASSASSASOLSAPSUKATCTTHHD 2220
Db 2161 PCEPEPDVTVLTSMLTDPSHITAEAKRRLARGSPPSLASSASSASOLSAPSUKATCTTHHD 2220
Qy 2221 SPDADLIEANLLWQEMGNIITRVESENKVVILDSFEPPLHAEGDEREISVAAEILRKSrk 2280
Db 2221 SPDADLIEANLLWQEMGNIITRVESENKVVILDSFEPPLHAEGDEREISVPAEILRKSrk 2280
Qy 2281 FPSALPIWARPDYNNPPLLESWKDPDYPPVHVHGCPLPFTKAPPIPPRRKRKTIVLTESNV 2340
Db 2281 FPSALPIWARPDYNNPPLLESWKDPDYPPVHVHGCPLPFTKAPPIPPRRKRKTIVLTESNV 2340
Qy 2341 SSALAEALATKTFGSSGSAVDSGTATLPDLASDGDGKSDVESVSMPPLEGPGDPL 2400
Db 2341 SSALAEALATKTFGSSGSAVDSGTATLPDLASDGDGKSDVESVSMPPLEGPGDPL 2400
Qy 2401 SDGSWSTVSEASDSDVVCSSMSYTWGTGALITPCCAEEESKLPINPLNSLLRHHNVYATT 2460
Db 2401 SDGSWSTVSEASDSDVVCSSMSYTWGTGALITPCCAEEETKLPINALNSLLRHHNVYATT 2460
Qy 2461 SRSASLRQKVTFRQLQVLDHRYDVLKEMKAKASTVKAKLLSIEBACKLTPPHSAKSP 2520
Db 2461 SRSASLRQKVTFRQLQVLDHRYDVLKEMKAKASTVKAKLLSIEBACKLTPPHSAKSP 2520
Qy 2521 GYGAKDVRLSSRAVNHIRSVMEDLLEDTEPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVRLSSRAVNHIRSVMEDLLEDTEPIDTTIMAKNEVFCVQPEKGRKPARLIV 2580
Qy 2581 FPDGLGVRCEKNALYDVVSTLPOAVMGSSSYGFOYSPQORVEFLVNTWKSKKCPMGFSYDT 2640
Db 2581 FPDGLGVRCEKNALYDVVSTLPOAVMGSSSYGFOYSPQORVEFLVNTWKSKKCPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVEESIYQCCDLAPBARQAIRSLTERLYTGGPLTNSKGQNGCYRRORA 2700
Db 2641 RCFDSTVTESDIRVEESIYQCCDLAPBARQAIRSLTERLYTGGPLTNSKGQNGCYRRORA 2700
Qy 2701 SGVLTTSCGNTLTCVLKATACRAAKLODCTMLVNGDDLVVICESAGTQEDAAALAFTE 2760
Db 2701 SGVLTTSCGNTLTCVLKASAAACRAAKLODCTMLVNGDDLVVICESAGTQEDANURVFTE 2760
Qy 2761 AMTRYSAAPPGPOPEYDLELITCSSNVSVVAHDASGRVYLLTRDPTTPLARAWEATAR 2820
Db 2761 AMTRYSAAPPGDLPOPEYDKELITCSSNVSVVAHDASGRVYLLTRDPTTPLARAWEATAR 2820
Qy 2821 HTPINSWLGNTIMVAPTLWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQ 2880
Db 2821 HTPVNSWLGNTIMVAPTLWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQ 2880
Qy 2881 IIERLHGLISATLHSSYSGEINRVASCLRKLGVPLRTWRHRAVSRAKLLSOGGRAATC 2940
Db 2881 IIORLHGLISATLHSSYSGEINRVASCLRKLGVPLRTWRHRAVSRAKLLSOGGRAATC 2940

```
QY 2941 GRYLFNNAVTKLUTIPAAASOLDLSGWFVAGYSGGDIYHSLSRARPRWFPLCLLLSV 3000
Db 2941 GRYLFNNAVTKLUTIPAAASOLDLSGWFVAGYSGGDIYHSLSRARPRWFPLCLLLSV 3000
QY 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010

RESULT 8
Q99AU2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q99AU2_9HEPC PRELIMINARY; PRT; 3010 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=chimera of HCV-BK;
RX MEDLINE=21534507; PubMed=11677216;
RA Thomson M., Nascimbene M., Gonzales S., Murthy K.K., Rehmann B.,
RA Liang T.J.;
RT "Emergence of a distinct pattern of viral mutations in chimpanzees
RT infected with a homogeneous inoculum of hepatitis C virus.";
RL Gastroenterology 121:1226-1233 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668 (1993).
DR EMBL; AF333324; AAK08509.1; -; mRNA.
DR PIR; A61196; A61196.
DR PIR; P0246; P0246.
DR PIR; P0804; P0804.
DR PIR; P50329; P50329.
DR SMR; Q99AU2; 1029-1657, 2008-2170.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008036; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
```

```
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV core; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327011 MW; 053B9A653B0AB335 CRC64;

Query Match 96.1%; Score 15392; DB 2; Length 3010;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2867; Conservative 71; Mismatches 72; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRKASERQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRKTSERQPRG 60
QY 61 RRQIPKARRPEGRAWAQPCYPMPLYGNEGGLWAGWLLSPRGRSPRGWPTDPRRRNLG 120
Db 61 RRQIPKARRPEGRWAQPCYPMPLYGNEGGLWAGWLLSPRGRSPRGWPTDPRRRNLG 120
QY 121 KVITDLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVITDLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIPASAYEVNRVSGIYHVTNDCNSSIVVEAADVTMHTPGVCVQEGNSSRCWV 240
Db 181 LLSCLTIPASAYEVNRVSGIYHVTNDCNSSIVVEAADVTMHTPGVCVQEGNSSRCWV 240
QY 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMTVGDLCSIFLVSQSLFTSPRRHET 300
Db 241 ALPTTLAARNSSIPTTIRRHVDLLVGAALCSAMVVDLCSGVFLVSQSLFTSPRYET 300
QY 301 VQDCNSIYPGHVSGHRMAMWMMNWSPTTALVVSQLLRIPQAVVDMVAGAHGVLGAGLA 360
Db 301 VQDCNSIYPGHVSGHRMAMWMMNWSPTTALVVSQLLRIPQAVVDMVAGAHGVLGAGLA 360
QY 361 YYSMVGWNAKVLIVALLFAGVDGTHVTGVRVAGHTTSGFTSLFSSGASQIQLVNTNGSW 420
Db 361 YYSMVGWNAKVLIVALLFAGVDGTHVTGVRVAGHTTSGFTSLFSSGASQIQLVNTNGSW 420
QY 421 HINRTALNCNDSLOTGFPAALFYAHKFNSSGCCPERMACRPIIDWFAQWGPITYTKPNSS 480
Db 421 HINRTALNCNDSLOTGFPAALFYAHKFNSSGCCPERMACRPIIDWFAQWGPITYTKPNSS 480
QY 481 DQRPYCHYAPRPGCVVPASQVCGPVYCFPTSPVVGTTDRSGVPTYSWGENETDVMMLN 540
Db 481 DQRPYCHYAPRPGCVVPASQVCGPVYCFPTSPVVGTTDRSGVPTYSWGENETDVMMLN 540
QY 541 NTRPPQGNWFGCTWMNSTGFTKCGPPCNIGGVGNRTLCPDTPRKHPHATYTKCGSG 600
Db 541 NTRPPQGNWFGCTWMNSTGFTKCGPPCNIGGVGNRTLCPDTPRKHPHATYTKCGSG 600
QY 601 PMLTPRCLVDYPRVLMHYPCITLNFSPKVRMYGVGVEHRLNAACNWTGRGRCNLEDDRS 660
Db 601 PMLTPRCLVDYPRVLMHYPCITLNFSPKVRMYGVGVEHRLNAACNWTGRGRCNLEDDRS 660
QY 661 ELSPLLLSTTEWQILPCAFTTLPALSTGLIHLHQNIVDVQYLYGVGSFAVFAIKWEYIL 720
Db 661 ELSPLLLSTTEWQILPCAFTTLPALSTGLIHLHQNIVDVQYLYGVGSFAVFAIKWEYIL 720
QY 721 LLFLLADARVCACLMWMLLIAQAEAALENLVLNAAASVAGAHGILSLFLVFFCAAWYIKG 780
```

Db 721 LFLFLLLADARVACACUWMLLIIAQAEATLENLVLNAASVAGAHGLLSFLVFFCAAWYIKG 780
QY 781 RLAPGAAYAFYGVWPLLLLLLLLALPRAYALDREMAASCGGAVLVLGLVPLTTPYKYKVELT 840
Db 781 RLVPGAAYALYGVWPLLLLLLLLALPRAYAMDREMAASCGGAVFVGLVLLTTPYKYKVELA 840
QY 841 RLILWLOVPIITRAEAHMOMVWVPLNVGRGRDAIILLTCVAPHELIIDITKLLLATILGLPM 900
Db 841 RLILWLOVPIITRAEAHLQWVWVPLNVGRGRDAIILLTCVAPHELIIDITKLLLATILGLPM 900
QY 901 VLQAGITRVPYFVRAQGLIRACMLVRKVAGGHYQVFMKLGALGTYYNHLTFLRDWA 960
Db 901 VLQAGITRVPYFVRAQGLIHACMLVRKVAGGHYQVFMKLGALGTYYNHLTFLRDWA 960
QY 961 HAGLRDLAVAVEPVVFSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HAGLRDLAVAVEPVVFSDMETKIIITWGADTAACGDIILGLPVSARRGKEIILGPADSLEG 1020
QY 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVUSTATQSFATCINGVCWT 1080
Db 1021 RGWRLLAPITAYSQOTRGLGCIITSLTGRDKNOVEGEVQVUSTATQSFATCINGVCWT 1080
QY 1081 VYHGAGSKTLAGKGPITQMYTNVDLVLVWQAPFGARSMTPCSGSSDLVLVTRHADVI 1140
Db 1081 VYHGAGSKTLAGKGPITQMYTNVDQDLVWQAPFGARSLTPCTCGSSDLVLVTRHADVI 1140
QY 1141 PVRRGDSRGSLLSPRPVSYLKGSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRGDSRGSLLSPRPVSYLKGSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVES 1200
QY 1201 METTWRSVPFTDNTSPPAVPQTFQVAHLHAPTGSGKSTKPAAYAAQGVKVLVLPNSVAA 1260
Db 1201 METTWRSVPFTDNTSPPAVPQSFQVAHLHAPTGSGKSTKPAAYAAQGVKVLVLPNSVAA 1260
QY 1261 TLGFGAYMSKAHGIDNIRGTVRTITGGSITYSTYKFLADGGCGSGGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAHGIDNIRGTVRTITGAPVYSTYKFLADGGCGSGGAYDIIICDECHS 1320
QY 1321 TDSSTTLIGTVLDQAETHAGALVLVATATPGSVTVPHNTEETGLSNNGEIPYKGAI 1380
Db 1321 TDSSTTLIGTVLDQAETHAGALVLVATATPGSVTVPHNTEEVALSNTGSEIPYKGAI 1380
QY 1381 PIEAIRGGRHLLIFCHSKKKCDELAACKLGLGNAYAYVGLDVSIVPTIGDVVVVATDAL 1440
Db 1381 PIEAIRGGRHLLIFCHSKKKCDELAACKLGLGNAYAYVGLDVSIVPTIGDVVVVATDAL 1440
QY 1441 MTGFTGDFSDVDCNTCVTQTVDFSLDPTFTTETTVPODAVSRSORRGRTGRSGIYR 1500
Db 1441 MTGFTGDFSDVDCNTCVTQTVDFSLDPTFTTETTVPODAVSRSORRGRTGRSGIYR 1500
QY 1501 FVTPGRRPSGMFSDSSVLCBCYDAGCAWYELTPAETSURLAYLNTPGLPVCQDHLFPWES 1560
Db 1501 FVTPGRRPSGMFSDSSVLCBCYDAGCAWYELTPAETSURLAYLNTPGLPVCQDHLFPWES 1560
QY 1561 VFTGLTHIDAHFLOSOTKOAGDNPPYLVAQATVCABAQAPPSWDMWKCLRKLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLOSOTKOAGDNPPYLVAQATVCARAQAPPSWDMWKCLRKLKPTLHG 1620
QY 1621 PTPLLYRLGAVQNEVILTHPIITKYINACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGVS 1680
Db 1621 PTPLLYRLGAVQNEVILTHPIITKYINACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGVS 1680
QY 1681 VIVGRILIISGKPAVPDREVLVQEFDEMBECCASQLPYIEQGMOLARQFKQKALGLQATAT 1740
Db 1681 VIVGRILIISGRPAIVPDRELLVQEFDEMBECCATHLPYIEQGMOLARQFKQKALGLQATAT 1740
QY 1741 KQAEAAAPVVESKWRALFTWAKHWNFIISGTOYLAGLSTLPCNPALIASIMAFATISIP 1800
Db 1741 KQAEAAAPVVESKWRALFTWAKHWNFIISGTOYLAGLSTLPCNPALIASIMAFATISIP 1800
QY 1801 LTTQNTLLFNILGSGWVAALAPPASAFYVAGIAGAAVGSIGLGVLDIILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGSGWVAALAPPASAFYVAGIAGAAVGSIGLGVLDIILAGYGAGVA 1860

Db 1801 LTTQSTLLFNILGSGWVAALAPPASAFYVAGIAGAAVGSIGLGVLDIILAGYGAGVA 1860
QY 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILLRHVGPGEGAVOMNRLLI 1920
Db 1861 GALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILLRHVGPGEGAVOMNRLLI 1920
QY 1921 AFASRGNHVSPTHYVVPESDAAARVTQILSLTITQLLKRHLQWINEDECSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHYVVPESDAAARVTQILSLTITQLLKRHLQWINEDECSTPCSGSWLRDV 1980
QY 1981 WDMTCTVLTDFTKWLQSKLLPRLPGVPFLSCORGYKGVWRGDGIMQITPCGCAQIAGHYK 2040
Db 1981 WDMTCTVLTDFTKWLQSKLLPQLPGVPFFSCORGYKGVWRGDGIMQITPCGCAQIAGHYK 2040
QY 2041 NGSMRIIVGPRTCSTNTHGTFPINAATTGCTPSPAPNYSRALWRVAABEYVEVTRVGDGFH 2100
Db 2041 NGSMRIIVGPRTCSTNTHGTFPINAATTGCTPSPAPNYSRALWRVAABEYVEVTRVGDGFH 2100
QY 2101 YVTGWTNDNVKPCQVPAPEFFTEVDGVRHLHRYAPACKPLLRREDVITFQVGLNQYLVSQ 2160
Db 2101 YVTGWTNDNVKPCQVPAPEFFSEVDGVRHLHRYAPACKPLLRREDVITFQVGLNQYLVSQ 2160
QY 2161 PCEPEPDVTVLTSMLTDPSHITAETAKRRLARGSPPSLASASSASQLSAPSLKATCTTHHD 2220
Db 2161 PCEPEPDVAVLTSMLTDPSHITAETAKRRLARGSPPSLASASSASQLSAPSLKATCTTHH 2220
QY 2221 SPDADLLEANLLWFOEMGNIITRVESENKVILDSFPLHAEGEREISVAAEILRKS 2280
Db 2221 SPDADLLEANLLWFOEMGNIITRVESENKVILDSFPLHAEGEREISVPAEILRKS 2280
QY 2281 FPSALPIWARPDYNPPLLESWKDPDYVPPVHVGCPLPPTKAPPIPPRRKRKTVVLTESNV 2340
Db 2281 FPAAMPPIWARPDYNPPLLESWKDPDYVPPVHVGCPLPPIKAPPIPPRRKRKTVVLTESV 2340
QY 2341 SSALAEALATKTFGSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPPLEGPGDD 2400
Db 2341 SSALAEALATKTFGSGSSAVDSGTATALPDQASDDGDKGSDVESYSSMPPLEGPGDD 2400
QY 2401 SDGWSSTVSEEAEDVVCSSMSYTWGALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
Db 2401 SDGWSSTVSEEAEDVVCSSMSYTWGALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
QY 2461 SRSASLRKKVYTFORLQVLDHYRDVLKEMAKASTVKAKLLSIEEACKLTTPPSAKSKF 2520
Db 2461 SRSAGLRKKVYTFORLQVLDHYRDVLKEMAKASTVKAKLLSVEEACKLTTPPSAKSKF 2520
QY 2521 GYGAKDVNRNLSRAVNHRSVWEDLLBDETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVNRNLSKAVNHIHSVMKOLLEDTVTPIDTTIMAKNEVFCVQPEKGRKPARLIV 2580
QY 2581 FPDILGVRVCEKALYDVVSTLPOAVMGSSYGFQYSPKQVSEFLVNTWKSKKCPMGFSYDT 2640
Db 2581 FPDILGVRVCEKALYDVVSTLPOVVMGSSYGFQYSPQVSEFLVNTWKSNNKPMGFSYDT 2640
QY 2641 RCFDSTVTESDIRVEESIOCCDLAPARQAIIRSLTERLYTGGPLTNSKGNCGYRRRA 2700
Db 2641 RCFDSTVTENDIRVEESIOCCDLAPARQAIKSLTERLYTGGPLTNSKGNCGYRRRA 2700
QY 2701 SGVLTTSCGNTLTCTCYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQOEAAAALRA 2760
Db 2701 SGVLTTSCGNTLTCTCYLKASACRAAKLQDCTMLVNGDDLVVICESAGTQOEAAAALRA 2760
QY 2761 AMTRYISAPGPPQPEYDLELITSCSSNVSAHADASKRVYVLTDPDPTPLARAWEATAR 2820
Db 2761 AMTRYISAPGPPQPEYDLELITSCSSNVSAHADASKRVYVLTDPDPTPLARAWEATAR 2820
QY 2821 HTPNSWLGNTIMVAPTLLWARMILMTHFFSILLAQOEKALDCQIYGACYSIEPLDLPQ 2880
Db 2821 HTPVNSWLGNTIMVAPTLLWARMILMTHFFSILLAQOEKALDCQIYGACYSIEPLDLPQ 2880
QY 2881 IIERLHGLSAPTLHSYSPGEINRVASCLURKUGVPPPLTRWHRARSVRAKLLSQGGRAATC 2940
Db 2881 IIERLHGLSAPSLHSYSPGEINRVASCLURKUGVPPPLRVWHRARSVRARLLSQGGRAATC 2940

QY 2941 GRYLFNVAWTKLTPIPAASOLDLSGWFVAGYSGGDIYHLSLRARPRWFPCLLLLSV 3000
 DB 2941 GKYLEFNVAWTKLTPIPAASOLDLSGWFVAGYSGGDIYHLSLRARPRWFPCLLLLSV 3000
 QY 3001 GVGIIYLLPNR 3010
 DB 3001 GVGIIYLLPNR 3010
 RESULT 9
 Q9QIX5_9HEPC
 ID Q9QIX5_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC Q9QIX5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polypeptide.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MD8-2;
 RA MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
 RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.-Y., Miyasaka Y.,
 RA Tazawa J.-i., Izumi N., Marumo F., Sato C.;
 RT "Time-related changes in full-length hepatitis C virus and hepatitis
 RT activity";
 RL Virology 263:244-253(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=92044457; PubMed=1658209;
 RA Oshina M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity";
 RL J. Gen. Virol. 72:2805-2809(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE
 RC MEDLINE=93224886; PubMed=8385694;
 RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
 RA Dusheiko G., Saeed A.A., Holmes E.C.;
 RT "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence
 RT diversity";
 RL J. Gen. Virol. 74:661-668(1993).
 DR EMBL; AF165060; AAD56195.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.
 DR PIR; PQ0254; PQ0254.
 DR PIR; PQ0804; PQ0804.
 DR PIR; PS0329; PS0329.
 DR HSP; Q9QIX5; 1CW.
 DR SMR; Q9QIX5; 1029-1657, 2008-2170, 2420-2949.
 DR MEROPS; S29.002; -;
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV_core.

DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV NS5b.
 DR InterPro; IPR001650; Helicase C.
 DR InterPro; IPR002518; Pept U39_HCV NS2.
 DR InterPro; IPR004109; Peptidase S29.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV capsid; 1.
 DR Pfam; PF01542; HCV core; 1.
 DR Pfam; PF01539; HCV env; 1.
 DR Pfam; PF01560; HCV NS1; 1.
 DR Pfam; PF01538; HCV NS2; 1.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00271; Helicase C; 1.
 DR Pfam; PF00998; RDRP_3; 1.
 DR SMART; SM00487; DEXDc; 1.
 KW Polyprotein.
 SQ SEQUENCE 3010 AA; 327298 MW; 8B99F1EBA6A50F56 CRC64;
 Query Match 96.1%; Score 15381; DB 2; Length 3010;
 Best Local Similarity 95.2%; Pred. No. 0;
 Matches 2866; Conservative 64; Mismatches 80; Indels 0; Gaps 0;
 QY 1 MSTNPKPQKTKNTNRRPQDVKFGGGOIVGGVYLLPRRGLPGLVRAATKASERSOPRG 60
 DB 1 MSTNPKPQKTKNTNRRPQDVKFGGGOIVGGVYLLPRRGLPGLVRAATKASERSOPRG 60
 QY 61 RRQPIKARRPEGRAWAQPCYPMPLXGNEGLWAGWLLSPRGRSPGWPDPTRRRSRNLG 120
 DB 61 RRQPIKARRPEGRWAQPCYPMPLXGNEGLWAGWLLSPRGRSPGWPDPTRRRSRNLG 120
 QY 121 KVDTLTTCGFDLWGVYPLVGLAPLGAARALAHGVRLVLEGVNATGNLPGCSFSIFLLA 180
 DB 121 KVDTLTTCGFDLWGVYPLVGLAPLGAARALAHGVRLVLEGVNATGNLPGCSFSIFLLA 180
 QY 181 LLSCLTIPASAYEVNVSIGIYHVTNDCNSSIYVEAADVIMHTPGVPCVQOEGNSSRCWV 240
 DB 181 LLSCLTIPASAYEVNVSIGIYHVTNDCNSSIYVEAADVIMHTPGVPCVQOEGNISRCWV 240
 QY 241 ALPTTLAARNASVPTTIRRHVDLLVGTAAFCSAMVYVDLCGSIPLVSQLFTSPRRHET 300
 DB 241 ALPTTLAARNISVPTTIRRHVDLLVGTAAFCSAMVYVDLCGSIPLVSQLFTSPRRHET 300
 QY 301 VQDCNCSIYPGHVSGHRMAMWMMNNSPTTALVVSQLLRIPQAVVDMVAGAHGVLGAGLA 360
 DB 301 VQDCNCSIYPGHVSGHRMAMWMMNNSPTTALVVSQLLRIPQAVVDMVAGAHGVLGAGLA 360
 QY 361 YYSWGVNWKVLIIVALLFAGVDGTHTTGVRAGHTTSGFTSLFSSGASQIKLVNTNGSW 420
 DB 361 YYSWGVNWKVLIIVALLFAGVDGTHTTGVRAGHTTSGFTSLFSSGASQIKLVNTNGSW 420
 QY 421 HINRTALNCNDSLOTGFFAALFYAHKFNSSGCCPERMASCRIPIWFAQGWGPITYTKPNSS 480
 DB 421 HINRTALNCNDSLOTGFFAALFYTHRNASGCPERMASCRIPIWFDQGWGPITYAEFNSP 480
 QY 481 DQRPYCHYAPRCGVVPAASQVCPVYCFPTSPVVTGTTDRSGVPTYSWGENETDVMLLN 540
 DB 481 DQRPYCHYAPRCGVVPAASQVCPVYCFPTSPVVTGTTDRSGVPTYSWGENETDVMLLN 540
 QY 541 NTRPPQGNWFCGCTWMNSTGTFTKCGGPPCNIGGVGNRTLICPTDPCFKHPEATVYKCGSG 600
 DB 541 NTRPPQGNWFCGCTWMNSTGTFTKCGGPPCNIGGVGNRTLICPTDPCFKHPEATVYKCGSG 600
 QY 601 PWLTFRCLVDYPRVRLWHYPCCTLANFSIFKVRMYGVGVHRLNAACNWTGRCNLEDNRDS 660
 DB 601 PWLTFRCLVDYPRVRLWHYPCCTVNFIFKVRMYGVGVHRLNAACNWTGRCNLEDNRDS 660

QY 661 ELSPLLLSTTEWQIILPCAFITTLPALSTGLIHLHONIVDVQVLYGVGSFAVFAIKWEYIL 720
Db 661 ELSPLLLSTTEWQIILPCSFITTLPALUSTGLIHLHONIVDVQVLYGVGSFAVFAIRWEYIL 720
QY 721 LFLFLLADARVCACILMMMLLIIAQAEALLENLVLNAASVAGAHGILSLVFFCAAWYIKG 780
Db 721 LFLFLLADARVCACILMMMLLIIAQAEALLENLVLNAASVAGAHGILSLVFFCAAWYIKG 780
QY 781 RLAPGAAYAFVGVWPELILLLLIALPPRAYALDREMAASCGGAVLVGLVFLTLSPYKYVFLT 840
Db 781 RLVPGAAYALYGVWPELILLLLIALPPRAYAMDREMAASCGGAVFVGLAFVFLTLSPYKYVFLA 840
QY 841 RLILWLQVFIITRAEAHMOMVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLALLGFLPM 900
Db 841 RLILWLQVFIITRAEAHLQVWVPLNVRGGRDAIILLCMCVHPELIFDITKLLALLGFLPM 900
QY 901 VLQAGITRVFVRAQGITRACMLVRKVAGGHYQVFMKLGALGTYYVNHLPRLDWA 960
Db 901 VLQAGITRVFVRAHGLIRACMLVRKVAGGHYQVMAFVKLAALTGTYYVDHLAPLQWA 960
QY 961 HAGLRDLAVAPVVPVFSAMETKVIITWGADTAACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HSLCLDLAVAPVVPVFSDMETKIIITWGADTAACGDIILGLPVSARRGREIILGPADSLEG 1020
QY 1021 QGWRLLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVSTATQSFATCINGVCWT 1080
Db 1021 QGWRLLAPITAYSQOTRGLLGCIITSLTGRDKNOVEGEVQVSTATQSFATCINGVCWT 1080
QY 1081 VYHGAGSKTLAGKGPITQMYTNVDLDLVGMQAPPGARSMTPCSCGSDLVLTVRHADVI 1140
Db 1081 VYHGAGSKTLAGKGPITQMYTNVDQDLVGMQAPPGARSMTPTCTCGSSDVLVTRHADVI 1140
QY 1141 PVRRGDRSGSLLSRPVSYLYKGSGGPLLCPSGHVGVFRAVCTRGVAKAVDPIPVES 1200
Db 1141 PVRRGDRGSGSLLSRPVSYLYKGSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDPIPVES 1200
QY 1201 METTWRSVPFTDNTSTPPAVPOTFOVAHLHAPGSGKSTKVPAAVAAQGVKVLVLPNSVAA 1260
Db 1201 METTWRSVPFTDNTSPPAVPOTFOVAHLHAPGSGKSTKVPAAVAAQGVKVLVLPNSVAA 1260
QY 1261 TLGFGAYMSKAGHDIPNIRGTVRTITTGGSITYSYGKFLADGCGSGGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKAGTDPNIRGTVRTITTGAPITYSYGKFLADGCGSGGAYDIIICDECHS 1320
QY 1321 TDSITLIGTGVLDQAETAGARLVLATATPPGSVTVPHNPTEERTGLSNNGSIPFYGKAI 1380
Db 1321 TDSITLIGTGVLDQAETAGARLVLATATPPGSVTVPHNPTEEVALSNTGEIPFYGKAI 1380
QY 1381 PIEAIKGGRHLLFCHSKKCCDELAACLGLGLNAVAYRGLDVSVIPPIGDVVVATDAL 1440
Db 1381 PIEVIKGGRHLLFCHSKKCCDELAACKLSALGUNAVAYRGLDVSVIPTSGDVVVVATDAL 1440
QY 1441 MTGFTGDFSVDICNTCVTQTVDFSLDPTFTTITTVPODAVSRSORGRTRGRSGIYR 1500
Db 1441 MTGYTGDFSVDICNTCVTQTVDFSLDPTFTTITTVPODAVSRSORGRTRGRGGIYR 1500
QY 1501 FVTPGRRPSGMFDSVSLCSCYDAGCAWYELTPAETSVRRLRAYLNTPGLPVOCODHLEFWES 1560
Db 1501 FVTPGRRPSGMFDSVSLCSCYDAGCAWYELTPAETSVRRLRAYLNTPGLPVCQDHLFEWES 1560
QY 1561 VFTGLTHIDAHFLSOTKQAGDNFPYLVAOATVCARAQAPPPSDOMMKCLIRLKPTLHG 1620
Db 1561 VFTGLTHIDAHFLSOTKQAGDNFPYLVAOATVCARAQAPPPSDOMMKCLIRLKPTLHG 1620
QY 1621 PIPLLYRLGAVONEVTLTHPIKYIMACMSADLEVVTSTWLVGGVLAALAYCLITTSV 1680
Db 1621 PIPLLYRLGAVONEVTLTHPIKYIMACMSADLEVVTSTWLVGGVLAALAYCLITTSV 1680
QY 1681 VIVGRILISGKPAVWPDRVLYQEFDEMEBCASQLPIEQGMOLAEQFKOKALGLLQAT 1740
Db 1681 VIVGRILISGKPAVWPDRVLYQEFDEMEBCASHLPIYIEQGMOLAEQFKOKALGLLQAT 1740

QY 1741 KOEAAAAPVVEVKWRALETETWAKHMWNFISGIIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KOEAAAAPVVEVKWRALETETWAKHMWNFISGIIQYLAGLSTLPGNPAIASLMAFTASITSP 1800
QY 1801 LTTQNTLLFNILGCVAAQALAPPSAASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGCVAAQALAPPSAASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVA 1860
QY 1861 GALVAFKVMSEVPSTEDLVNLLPAILSPCALVGVVCAAILRHHVGPGEAGVOMMRLLI 1920
Db 1861 GALVAFKVMSEVPSTEDLVNLLPAILSPCALVGVVCAAILRHHVGPGEAGVOMMRLLI 1920
QY 1921 AFASRGNHVPTHVVPESDAAARVTQILSSILITQLLKRHLQWINECDSTPCSCSWLRDV 1980
Db 1921 AFASRGNHVPTHVVPESDAAARVTQILSSILITQLLKRHLQWINECDSTPCSCSWLRDV 1980
QY 1981 WDMICTVLTFKTLWQSKIMPRLPGVPFLSCQRGYKGVMRGDIQMOTTCPCGAQITGHVK 2040
Db 1981 WDMICTVLTFKTLWQSKIMPRLPGVPFLSCQRGYKGVMRGDIQMOTTCPCGAQITGHVK 2040
QY 2041 NGSMRIIVGPBKTCSNTHGTFPINAATYTGCTPSPAPNYSRALWVAEEYEVTVRGDFH 2100
Db 2041 NGSMRIIVGPBKTCSNTHGTFPINAATYTGCTPSPAPNYSRALWVAEEYEVTVRGDFH 2100
QY 2101 YVTGMTTDNVKPCQVPAPEFFTEVDGVRHLRVAPACKPLLRDVTTFQVGLNQVLVGSOL 2160
Db 2101 YVTGMTTDNVKPCQVPAPEFFTEVDGVRHLRVAPACKPLLRDVTTFQVGLNQVLVGSOL 2160
QY 2161 PCEPEPDVTLTMSLTDPHSITAEAKRRLARGSPPSLASSASSQSLSAPSLSKATCTTHHD 2220
Db 2161 PCEPEPDVTLTMSLTDPHSITAEAKRRLARGSPPSLASSASSQSLSAPSLSKATCTTHHD 2220
QY 2221 SPDADLLEANLLWROENGNIITRVESENKVVILDSFPLHAEGDEREISVAAEILRSRK 2280
Db 2221 SPDADLLEANLLWROENGNIITRVESENKVVILDSFPLHAEGDEREISVAAEILRSRK 2280
QY 2281 FPSALPIWARPDPNPPLESKWDPDYVPVHVHGCPLPPTKAPPIPPPRKRKTVVLTESNV 2340
Db 2281 FPSALPIWARPDPNPPLESKWDPDYVPVHVHGCPLPPTKAPPIPPPRKRKTVVLTESNV 2340
QY 2341 SSALAEALATKTFGSGSSAVDSGTATAPDQASDDGDKESVSYSSMPPLEGSPGDPDL 2400
Db 2341 SSALAEALATKTFGSGSSAVDSGTATAPDQASDDGDKESVSYSSMPPLEGSPGDPDL 2400
QY 2401 SDGWSVTSVEBASDDVVCSSSYTWTGALITPCAAESKLPINPLNSLLRHHNNVYATT 2460
Db 2401 SDGWSVTSVEBASDDVVCSSSYTWTGALITPCAAESKLPINPLNSLLRHHNNVYATT 2460
QY 2461 SRSASLRQKVTFFORLOVLDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSF 2520
Db 2461 SRSASLRQKVTFFORLOVLDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSF 2520
QY 2521 GYGAKDVNRNLSRRVNHRSVWEDLLEDETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVNRNLSRRVNHRSVWEDLLEDETETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580
QY 2581 FPDILGVRVCEKMAIYDVVSTLPQAVMGSSYCFQYSPKQVSEFLVNTWKSCKCPMGFSYDT 2640
Db 2581 FPDILGVRVCEKMAIYDVVSTLPQAVMGSSYCFQYSPKQVSEFLVNTWKSCKCPMGFSYDT 2640
QY 2641 RCFDSTVTESDIRVEESIQCCDLAPARQAIRSILTERLYIGGPLTNSKGQVYRRCRA 2700
Db 2641 RCFDSTVTESDIRVEESIQCCDLAPARQAIRSILTERLYIGGPLTNSKGQVYRRCRA 2700
QY 2701 SGVLTTSNGTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICSSAGTQEDAAALRAFTE 2760
Db 2701 SGVLTTSNGTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICSSAGTQEDAAALRAFTE 2760
QY 2761 AMTRYSAPPGDPPOPEYDLELITSCSSNVSVAHADSGKRVYLTDPPTPLARAWEATAR 2820
Db 2761 AMTRYSAPPGDPPOPEYDLELITSCSSNVSVAHADSGKRVYLTDPPTPLARAWEATAR 2820
QY 2821 HTPINSWMLGNIIMVAPTILWARMILMTHPFFSILLAQEQLKALDCQIYGACYSIEPLDLPO 2880

```
DB 2821 HTFVNSWLNIIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEBPLDLPQ 2880
QY 2881 IIERHLGLSAFTLHSSYSPGEINRVASCKRLGVPPLRTWHRARSVRKLLSQGGRAATC 2940
DB 2881 IIERHLGLSAFSLHSSYSPGEINRVASCKRLGVPPLRVHRARSVRKLLSQGGRAATC 2940
QY 2941 GRYLFNNAVTRTKLTPIPAASOLDLSGMFVAGYSGDDIYHLSLRARPWFPLCLLLSV 3000
DB 2941 GRYLFNNAVTRTKLTPIPAASRLDLSGMFVAGYSGDDIYHLSLRARPWFPMWCLLLSV 3000
QY 3001 GVGIYLLPNR 3010
DB 3001 GVGIYLLPNR 3010

RESULT 10
Q9J3H7_9HEPC PRELIMINARY; PRT; 3010 AA.
ID Q9J3H7;
AC Q9J3H7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD15;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshina M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.;
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL: AF207756; AAF65946.1; -; Genomic_RNA.
DR PIR: A61196; A61196.
DR PIR: P0246; P0246.
DR PIR: P0804; P0804.
DR PIR: PS0329; PS0329.
DR HSSP: Q8JYS1; 1CWX.
DR SMR: Q9J3H7; 1029-1657, 2008-2170, 2420-2949.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV_core.
```

```
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV NS5b.
DR InterPro: IPR002518; PpE_U39; HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00398; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327370 MW; D8653F7317FFA106 CRC64;

Query Match 96.1%; Score 15378; DB 2; Length 3010;
Best Local Similarity 95.0%; Pred.No. 0;
Matches 2860; Conservative 72; Mismatches 78; Indels 0; Gaps 0;

QY 1 MSTNPKPQRTKNTNRRPQDVKFGGGQIVGGVYLLPRRGLRVATRKASERSQPRG 60
DB 1 MSTNPKPQRTKNTNRRPQDVKFGGGQIVGGVYLLPRRGLRVATRKASERSQPRG 60
QY 61 RRQIPKARRPEGRAWQPGYPMPLVNGEGLWAGWLLSPGSRPSPGTPDPRRSNLG 120
DB 61 RRQIPKARRPEGRWTAQPGYPMPLVNGEGLWAGWLLSPGSRPSPGTPDPRRSNLG 120
QY 121 KVTDLTTCGFADLGMVYPLVGLGPAALAHGVLEGVNATGNLPGCSFIFLLA 180
DB 121 KVTDLTTCGFADLGMVYPLVGLGPAALAHGVLEGVNATGNLPGCSFIFLLA 180
QY 181 LLSCLTIPASAYEVRNVSGIYHVTNDCNSSIYVEAADVIMHTFPGCVQEGNSRCWV 240
DB 181 LLSCLTIPASAYEVRNVSGIYHVTNDCNSSIYVEAADVIMHTFPGCVQEGNSRCWV 240
QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVDLCCGSIPLVSLFTSPRHET 300
DB 241 ALTPTLAARNISPTTIRRHVDLLVGTAAFCSAMVVDLCCGSIPLVSLFTSPRHET 300
QY 301 VQDCNSIYDGHVSGHRMAMDMNNSPTTALVVSQLLRIPQAVDMVAGAHGVLGALA 360
DB 301 VQDCNSIYDGHVSGHRMAMDMNNSPTTALVVSQLLRIPQAVDMVAGAHGVLGALA 360
QY 361 YYSVGNWAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW 420
DB 361 YYSVGNWAKVLIVALLFAGVDGETHTTGRVAGHTTSGFTSLFSSGASQIKLVNTNGSW 420
QY 421 HINRTALNCNDSLQTFGAALFYAHKFNSSGCPERMASCRPIDMFAQGWGPIITTKNSS 480
DB 421 HINRTALNCNDSLQTFGAALFYTHRFNSGCPERMASCRPIDMFAQGWGPIITTKNSS 480
QY 481 DQRPYCHYAPRCVVPASQVCGPVYCFPTSPVAVCTTDRSGVPTYSWGENETDVLN 540
DB 481 DQRPYCHYAPRCVVPASQVCGPVYCFPTSPVAVCTTDRSGVPTYSWGENETDVLN 540
QY 541 NTRPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLCPTCFRKHGPATYTKCGSG 600
DB 541 NTRPQGNWFGCTWMNSTGFTKTCGGPPCNIGGVGNRTLCPTCFRKHGPATYTKCGSG 600
QY 601 PWLTFRCLVDVYPYRLMHPCTLNFSIFKVRMYVGVGVEHRLNAACNWTGRGECNLEDRDRS 660
DB 601 PWLTFRCLVDVYPYRLMHPCTLNFSIFKVRMYVGVGVEHRLNAACNWTGRGECNLEDRDRS 660
QY 661 ELSPLLSTTETWQILPCAFTTLPALSTGLIHLHONIVDVQVLYGVGSFAFVFAIKWEYIL 720
```

[illegible]

QY 2881 IIERLHLSAFTLHYSVPEINRVASCLRLKLVPPPLRTWRHRSVRKLLSQGGAATC 2940
 DB 2881 IIERLHLSAFTLHYSVPEINRVASCLRLKLVPPPLRVHRARSVRKLLSQGGAATC 2940
 QY 2941 GRYLENWVATKLTPIPAASOLDLSGWFVAGYSGDDIYHLSRAPRPFPLCLLLLSV 3000
 DB 2941 GRYLENWVATKLTPIPAASOLDLSGWFVAGYSGDDIYHLSRAPRPFPLCLLLLSV 3000
 QY 3001 GVGIVLLPNR 3010
 DB 3001 GVGIVLLPNR 3010
 RESULT 11
 POLG HCVAJ STANDARD; PRT; 3009 AA.
 AC P26622; P89966; Q81755;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Genome polyprotein [Contains: Core protein p19 (Capsid protein C)
 DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein
 DE E2 (NS1) (gp68) (gp70); p7; Protease NS2-3 (EC 3.4.22.-) (p23); Serine
 DE protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.-)
 DE (Hepacivirin) (NS3P) (p70); Nonstructural protein 4A (NS4A) (p8);
 DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
 DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
 OS Hepatitis C virus (isolate Japanese) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OC NCBI_TaxID=11116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
 RX MEDLINE-91088550; PubMed-2175903;
 RA Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
 RA Sugimura T., Shimotohno K.;
 RT "Molecular cloning of the human hepatitis C virus genome from Japanese
 RT patients with non-A, non-B hepatitis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
 RN [2]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE-91192160; PubMed-1849488; DOI=10.1016/0014-5793(91)80322-T;
 RA Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraio K.,
 RA Ohkoshi S., Shimotohno K.;
 RT "Molecular structure of the Japanese hepatitis C viral genome.";
 RL FEBS Lett. 280:325-328(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
 RA Tanaka T.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE [GENOMIC RNA] OF 722-1907, IDENTIFICATION OF BOTH
 RP PROTEASES, AND MUTAGENESIS OF CYS-921; HIS-931; HIS-951; GLU-971;
 RP GLU-979; CYS-992; GLU-1008; GLU-1057; HIS-1082; ASP-1106; CYS-1122;
 RP CYS-1124; HIS-1135; SER-1164; CYS-1170; HIS-1174; CYS-1184; GLU-1198;
 RP GLU-1201; HIS-1226 AND HIS-1228.
 RX MEDLINE-9323208; PubMed-8392606;
 RA Hijikata M., Mizushima H., Akagi T., Mori S., Kakiuchi N., Kato N.,
 RA Tanaka T., Kimura K., Shimotohno K.;
 RT "Two distinct proteinase activities required for the processing of a
 RT putative nonstructural precursor protein of hepatitis C virus.";
 RL J. Virol. 67:4665-4675(1993).
 RN [5]
 RP INHIBITION OF HEPATITIS B VIRUS GENE EXPRESSION.
 RX PubMed-8396658;
 RA Shih C.-M., Lo S.-J., Miyamura T., Chen S.-Y., Lee Y.-H.W.;
 RT "Suppression of hepatitis B virus expression and replication by
 RT hepatitis C virus core protein in Huh-7 cells.";
 RL J. Virol. 67:5823-5832(1993).
 RN [6]
 RP PROCESSING OF POLYPROTEIN, MUTAGENESIS OF HIS-1082; ASP-1106 AND
 RP SER-1164, AND IDENTIFICATION OF THE SERINE PROTEASE.

RX PubMed-8389908;
 RA Bartenschlager R., Ahlborn-Laake L., Mous J., Jacobsen H.;
 RT "Nonstructural protein 3 of the hepatitis C virus encodes a serine-
 RT type proteinase required for cleavage at the NS3/4 and NS4/5
 RT junctions.";
 RL J. Virol. 67:3835-3844(1993).
 RN [7]
 RP PHOSPHORYLATION OF NS5A.
 RX PubMed-7999043;
 RA Kaneko T., Tanji Y., Satoh S., Hijikata M., Asabe S., Kimura K.,
 RA Shimotohno K.;
 RT "Production of two phosphoproteins from the NS5A region of the
 RT hepatitis C viral genome.";
 RL Biochem. Biophys. Res. Commun. 205:320-326(1994).
 RN [8]
 RP MUTAGENESIS OF SER-2193; SER-2196; SER-2199; SER-2200; SER-2201;
 RP SER-2203; SER-2206; SER-2209 AND SER-2220.
 RX PubMed-7769656;
 RA Tanji Y., Kaneko T., Satoh S., Shimotohno K.;
 RT "Phosphorylation of hepatitis C virus-encoded nonstructural protein
 RT NS5A.";
 RL J. Virol. 69:3980-3986(1995).
 RN [9]
 RP FUNCTION OF NS5A.
 RX MEDLINE-96127954; PubMed-8531962; DOI=10.1056/NEJM199601113340203;
 RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
 RA Yamamoto C., Ogura Y., Izumi N., Maruno F., Sato C.;
 RT "Mutations in the nonstructural protein 5A gene and response to
 RT interferon in patients with chronic hepatitis C virus 1b infection.";
 RL N. Engl. J. Med. 334:77-81(1996).
 RN [10]
 RP FUNCTION OF NS5A.
 RX PubMed-9343247;
 RA Kato N., Lan K.H., Ono-Nita S.K., Shiratori Y., Omata M.;
 RT "Hepatitis C virus nonstructural region 5A protein is a potent
 RT transcriptional activator.";
 RL J. Virol. 71:8856-8859(1997).
 RN [11]
 RP ZINC-BINDING REGION OF SERINE PROTEASE NS3, COFACTOR, AND MUTAGENESIS
 RP OF CYS-1041; CYS-1072; CYS-1077; HIS-1082; HIS-1135; SER-1164;
 RP HIS-1174; CYS-1184; HIS-1226 AND HIS-1228.
 RX PubMed-9060645;
 RA Stempiak M., Hostomska Z., Nides B.R., Hostomsky Z.;
 RT "The NS3 proteinase domain of hepatitis C virus is a zinc-containing
 RT enzyme.";
 RL J. Virol. 71:2881-2886(1997).
 RN [12]
 RP SUBCELLULAR LOCATION OF CORE PROTEIN.
 RX PubMed-9621068;
 RA Yasui K., Wakita T., Tsukiyama-Kohara K., Funahashi S.-I.,
 RA Ichikawa M., Kajita T., Moradpour D., Wands J.R., Kohara M.;
 RT "The native form and maturation process of hepatitis C virus core
 RT protein.";
 RL J. Virol. 72:6048-6055(1998).
 RN [13]
 RP ISDR REGION.
 RX PubMed-9525599;
 RA Pawlotsky J.-M., Germanidis G., Neumann A.U., Pellerin M.,
 RA Franaig P.-O., Dhumeaux D.;
 RT "Interferon resistance of hepatitis C virus genotype 1b: relationship
 RT to nonstructural 5A gene quasispecies mutations.";
 RL J. Virol. 72:2795-2805(1998).
 RN [14]
 RP OLIGOMERIZATION OF NS5B.
 RX PubMed-11907226; DOI=10.1128/JVI.76.8.3865-3872-2002;
 RA Wang Q.M., Hockman M.A., Staschke K., Johnson R.B., Case K.A., Lu J.,
 RA Parsons S., Zhang F., Rathnachalam K., Kirkgaard K., Colacino J.M.;
 RT "Oligomerization and cooperative RNA synthesis activity of hepatitis C
 RT virus RNA-dependent RNA polymerase.";
 RL J. Virol. 76:3865-3872(2002).
 RN [15]
 RP INHIBITION OF P7 BY AMANTADINE.
 RX MEDLINE=22448611; PubMed=12560074; DOI=10.1016/S0014-5793(02)03851-6;


```
QY 1082 YHAGSKTLGPKGPIQMTYNTVDLDLVGWQAPPGARSMTPCSGSSDLYLVTRHADVIP 1141
Db 1081 YHAGSKTLGPKGPIQMTYNTVDLDLVGWQAPPGARSMTPCSGSSDLYLVTRHADVIP 1140
QY 1142 VRRRGSRSGLSPRPVSLKSSGGPLCPGSHVVGVPFAAVCTRGVAKAVDFIPVESM 1201
Db 1141 VRRRGSRSGLSPRPVSLKSSGGPLCPGSHVVGIPFAAVCTRGVAKAVDFIPVESM 1200
QY 1202 ETTMRSPTFTNSTPPAVPQTQVAHLHAPTSGSKSTKVPAAAYAGYKVLVLPNSVAAT 1261
Db 1201 ETTMRSPTFTNSTPPAVPQTQVAHLHAPTSGSKSTKVPAAAYAGYKVLVLPNSVAAT 1260
QY 1262 LGFGAYMSKAHGDIPNIRGTFTTGGGSIYTSYOKFLADGGCSGGAYDIIICDECHST 1321
Db 1261 LGFGAYMSKAHGIERNIRGTFTTGGGPIYTSYCKFLADGGCSGGAYDIIICDECHST 1320
QY 1322 DSTTILIGTVDLDOAETAGARLVVLTATPPGCVTPPHNIEEIGLSNNGEIPFYGAIP 1381
Db 1321 DSTTILIGTVDLDOAETAGARLVVLTATPPGCVITVPHNIEEVALNSNTGEIPFYGAIP 1380
QY 1382 ISAIKGRHLIFCHSKKKDELAALKTLGLNNAVAYRGLDVSVIPIGDVVVVAATDALM 1441
Db 1381 ISAIKGRHLIFCHSKKKDELAALKTLGLNNAVAYRGLDVSVIPTSGDVVVVAATDALM 1440
QY 1442 TGFSGDFSDVIDCNTCVTQTVDFSLDPTFTTETTPQDAVRSORGRGTGRSGIYRF 1501
Db 1441 TGFSGDFSDVIDCNTCVTQTVDFSLDPTFTTETTLTPQDAVRSQAQRGTGRSGIYRF 1500
QY 1502 VTPGERPSGMFDSVLCBVDACGANVELTPAETSVRLRAYLNTGRLPVCOOHLFEWESV 1561
Db 1501 VTPGERPSGMFDSVLCBVDACGANVELTPAETSVRLRAYLNTGRLPVCOOHLFEWESV 1560
QY 1562 FTGLTHIDAHFLSQTQAGDNFPYLVAYQATVCARAQPPSDQWQKCLIRLKPILHGP 1621
Db 1561 FTGLTHIDAHFLSQTQAGDNLPYLVAQATVCARAQPPSDQWQKCLIRLKPILHGP 1620
QY 1622 TPLLYRLGAVQNEVILTHPTIKYIMACMSADLEVSTSTWLVGGVLAALAAAYCLTTGSVV 1681
Db 1621 TPLLYRLGAVQNEVILTHPTIKYIMACMSADLEVSTSTWLVGGVLAALAAAYCLTTGSVV 1680
QY 1682 IVGRILISGKPAVPDREVLVYQFDEMEBCASQIPYIEQGMQALAEQKQKALGLQATK 1741
Db 1681 IVGRILISGRPAVIPDREVLVYQFDEMEBCASHLPYIEQGMQALAEQKQKALGLQATK 1740
QY 1742 QAEEAAPVVEKRALETFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPL 1801
Db 1741 QAEEAAPVVEKRALEVFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPL 1800
QY 1802 TTQNTLLFNILGGVAAQAAPPASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVAG 1861
Db 1801 TTQNTLLFNILGGVAAQAAPPASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVAG 1860
QY 1862 ALVAFKVMGSEVPSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEAGVOMNRLIA 1921
Db 1861 ALVAFKVMGSEMPSTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEAGVOMNRLIA 1920
QY 1922 FASRGNHVSPTHYPESDAARVTOILSSLTITQLLKLHQTWINECDSTPCSGSLRDVW 1981
Db 1921 FASRGNHVSPTHYPESDAARVTOILSSLTITQLLKLHQTWINECDSTPCSGSLKDVW 1980
QY 1982 DWICTVLTDFKTLQSKLLPRLPCVPFLSCORGKYGVMRGDGINQTTCPGCAQIAGHVKN 2041
Db 1981 DWICTVLSDFKTLQSKLLPRLPGLPFLSCORGKYGVMRGDGINQTTCPGCAQIIGHVKN 2040
QY 2042 GSMRIVGERTCSNTWHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYEVETRVGDFHY 2101
Db 2041 GSMRIVGPKTCSNTWHGTFPINAYTTGCTPSPAPNYSRALWRVAABEYEVETRVGDFHY 2100
QY 2102 VTGMTNTNVKPCQVPAPEPTEVDGVLRLHRYAPACKPLIBEDYTFQVGLNOYLAVGSQLP 2161
Db 2101 VTGMTNTNVKPCQVPAPEPTEVDGVLRLHRYAPACKPLIBEEYVFGVGLNQYLAVGSQLP 2160
```

```
QY 2162 CEPEPDVTVLSMLTDPHSHTAETAKRLARGSPPSLASSSSASQLSAPSLKATCTTHDS 2221
Db 2161 CEPEPDVAVLVLSMLTDPHSHTAETAKRLARGSPPSLASSSSASQLSAPSLKATCTTHDS 2220
QY 2222 PDADLIEANLLWRQEMGNITRVESENKVVILDSFPEPLHAEGDEREISVAAEILRKRKF 2281
Db 2221 PDADLIEANLLWRQEMGNITRVESENKVVILDSFDPRAVEDEREISVPABILRKRKF 2280
QY 2282 PSALPIWARPDPYNPILLESWKDPDVPVPHVHCPLPPTKAPPIPPRRKRTVVLTESNVS 2341
Db 2281 PPALPIWARPDPYNPILLESWKDPDVPVPHVHCPLPSTKAPPIPPRRKRTVVLTESVVS 2340
QY 2342 SALAELATKTFSSGSSAVDSGTATATLPDASDDGDKGSDVESYSMPLEGEPPGPDLS 2401
Db 2341 SALAELATKTFSSGSSAVDSGTATGPPDOASDDGDKGSDVESYSMPLEGEPPGPDLS 2400
QY 2402 DGSWSTVSEASEDVVCCSMSTWTGALITPCAABESKLPINPLNSLLRHNNVATTS 2461
Db 2401 DGSWSTVSEAGEDVVCCSMSTWTGALITPCAABESKLPINPLNSLLRHHSMTSTTS 2460
QY 2462 RSASLRQKVTFRLOVLDHYRDVLKEMKAKASTVKALLSIEEACKLTPPHSAKSKFG 2521
Db 2461 RSASLRQKVTFRLOVLDHYRDVLKEMKAKASTVKARLLSIEEACKLTPPHSAKSKFG 2520
QY 2522 YGAKDVRLSSRAVNHIRSWEDELTETPIDTTIMAKSEVFCVQPEKGRKPARLIVF 2581
Db 2521 YGAKDVRSUSSRAVNHIRSWEDELTETPIDTTIMAKNEVFCVQPEKGRKPARLIVF 2580
QY 2582 PDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQVSPQRVEFLVNTWKSCKCPMGFSYDTR 2641
Db 2581 PDLGVRVCEKMALYDVVSTLPQAVMGPSYGFQVSPQRVEFLVNTWKSCKCPMGFSYDTR 2640
QY 2642 CFDSVTESDRIVVEESIYOCCLAPPEARQAIIRSLTERLYIGGPLTNSKGONCGYRRCRAS 2701
Db 2641 CFDSVTENDIRTEESIYOCCLAPPEARQAIIRSLTERLYVGGPLTNSKGONCGYRRCRAS 2700
QY 2702 GVLTTCGNTLTCYLKATAACRAAKLQDCTMLVNGDLDLVVICESAGTQEDAAALRAFTEA 2761
Db 2701 GVLTTCGNTLTCYLKATAACRAAKLQDCTMLVNGDLDLVVICESAGTQEDAAALRAFTEA 2760
QY 2762 MTRYSAPPDPPPOPEYDLELITSCSSNVSAHDASGRVYVLTTRDPTTPLARAAMETARH 2821
Db 2761 MTRYSAPPDPPPOPEYDLELITSCSSNVSAHDASGRVYVLTTRDPTTPLARAAMETVRH 2820
QY 2822 TPINSWLGNIIMVAPTLWARMILMTHFFSILLAOBQLEKALDCQIYGACYSIEPLDLPOI 2881
Db 2821 TPVNSWLGNIIMVAPTLWARMILMTHFFSILLAOBQLEKALDCQIYGACYSIEPLDLPOI 2880
QY 2882 IERLHGLSAPTLHSYSPGEINRVASCLKLGVPPLRTWHRARSVRAKLLSOGGRAATCG 2941
Db 2881 IERLHGLSAPSLHSYSPGEINRVASCLKLGVPPLRVHRARSVRAKLLSOGGRAATCG 2940
QY 2942 RYLFPNVAVRTKLTLPIPAASQLDLSGFWVAGYSGGDIYHLSRARPRPFWPLCLLLLSVG 3001
Db 2941 KYLFPNVAVRTKLTLPIPAASQLDLSGFWVAGYNGGDIYHLSRARPRPFWPLCLLLLSVG 3000
QY 3002 VGIYLLPNR 3010
Db 3001 VGIYLLPNR 3009

RESULT 12
Q9J3H5_9HEPC
ID Q9J3H5_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3H5_
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
```

RV NUCLEOTIDE SEQUENCE.
RC STRAIN=MD17;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207758; AAF65948.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWX.
DR SMR; Q9J3H5; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S23.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_Psvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF0271; Helicase_C; 1.
DR Pfam; PF00998; RGRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326806 MW; 9FE3D1B93B7AA4B CRC64;

Query Match 95.9%; Score 15358; DB 2; Length 3010;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 2874; Conservative 57; Mismatches 78; Indels 2; Gaps 2;

QY 1 MSTNPKPQKTKRNTNRPDQVKKFGGQIVGGVYLLPRRGPRLGVRATRKASERSQPRG 60
DB 1 MSTNPKPQKTKRNTNRPDQVKKFGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

QY 61 RQPIPKARRPEGRAWAQPGYPWPVLYGNEGLGWAGWLLSPGRSRPSWGPTDPRRSRLG 120
DB 61 RQPIPKARQPEGRAWAQPGYPWPVLYGNEGMWAGWLLSPGRSRPSWGPTDPRRSRLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEQVNVYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEQVNVYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTIPASAYEVRNVSGIYHVTNDCSNSSIVYEAADVIMHTPGCVPCVOEGNSSRCWV 240
DB 181 LLSCLTIPASAYEVRNVSGAYHVHTNDCSNSSIVYETADMIMHTPGCVPCVREGNHSRCWA 240
QY 241 ALPTLAARNASVPTTIRRHVDLLVCTAAFCSMYVGDLCGSIFLYSOLFTSPRHHET 300
DB 241 ALPTLAARNASVPTTIRRHVDLLVGAAPFCSMYVGDLCGSVFLYSOLFTSPRHHET 300
QY 301 VQDCNCSIYPGHYSGHRMAMDMNNSPTTALVVSQLLRIPQAVVDMVAGAHWVLAGLA 360
DB 301 VQDCNCSIYPGHYSGHVMAMDMNNSPTTALVVSQLLRIPQAVVDMVAGAHWVLAGLA 360
QY 361 YYSVMGNWAKVLIIVALLFAGVDGE-THTTGVRAGHTTSGFTSLFSSGASQKIQIIVNTNGS 419
DB 361 YYSVMGNWAKVLIIVMLLFAGVDGTGLVSGVSGHTTSLTSLFRVGASQRIQLVNTNGS 420
QY 420 WHINRTALNCNDSLOTGFFAALFYAHKFNSSGCGPERNASCRPIDWFAQGWPIITYTPENS 479
DB 421 WHINRTALNCNDSLOTGFLAALFYTNKFNSSGCGSERLNSCRPIENFTQGWGPIITYDEGG 480
QY 480 SDORPYCWHYAPPCGVVPASQVCGPVYCTPSPVVVGTDRSGVPTYSWGENETDVMLL 539
DB 481 SDQPCYCHYAPQCQGIIVPASEVCGPVYCTPSPVVVGTDRDLGVPTYSWGENETDVMLL 540
QY 540 NNTRPQGNWFGCTWMNSTGFTKCGPPCNIGGVGNRTLICPTDCFRKHPEATYTKGS 599
DB 541 NNTRPQGNWFGCTWMNSTGFTKCGPPCNIGGVGNR-LVCPDTCFRKHPEATYTKGS 599
QY 600 GPMLTPRCLVDYPRYLWHYPTCLNFSIFKVRMYVGGVEHRLNAACNTRGRERCMLEDRD 659
DB 600 GPMLTPRCLVDYPRYLWHYPTCLNFSIFKVRMYVGGVEHRLNSAACNTRGRERCMLEDRD 659
QY 660 SELSPLLSLTTEWQILPCAFITLIPALSTGLIHQNIIVDVQYLVGVSFAFVKWEYI 719
DB 660 SELSPLLSLTTEWQILPCSTFTLLPALSTGLIHQNIIVDVQYLVGVSFAFVKWEYI 719
QY 720 LLLFLLIADARVCACLMMMLLIAQEAALLENVLNAAASVAGAHGILSFLVFCAAWYIK 779
DB 720 LLLFLLIADARVCACLMMMLLKQAEALLENVLNAAASVAGAHGILSFLVFCAAWYIK 779
QY 780 GRLAPGAAYAFYGVWPLLLLLLPPRAYALDRMAASCGGAVLVGLVFLTSLPYKVF 839
DB 780 GRLAPGVAYAFYGVWPLLLLLLPPRAYAMDRMAASCGGAVLVGLVFLTSLPHYKVFL 839
QY 840 TRLIWLQYFITRAEAHMVWVPLNVRRGGRDAIILLTCAVHPBELIFDITKLLAILGLPL 899
DB 840 TRIIWLQYFITRAEAHLQVWVPLNVRRGGRDAIILLTCAVHPBELIFDITKLLAILGLPL 899
QY 900 MVLOAGITRVYPYFVRAOGLIRACMLVRKAGGHYVQVMFKLGALTGTYYVNHITPLRDW 959
DB 900 MVLOAGITRVYPYFVRAOGLIRACMLVRKAGGHYVQVMALMKLAALTGTYYVNHITPLRDW 959
QY 960 AHAGIRDLAVAVEPVFSAMETKVIITWGAADTAACGDIILGLPVSARGKEIFLGPADSLE 1019
DB 960 AHTGLRDLAVAVEPVFSAMETKIITWGAADTAACGDIISGLPVSARGREILLGPADSLE 1019
QY 1020 GQGRLLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVVSATQSFATCINGVCW 1079
DB 1020 GQGRLLAPITAYSQOTRGVLGCIITSLTGRDKNOVEGEVQVVSATQSFATCINGVCW 1079
QY 1080 TVYHGAGSKTLIAGPKGPIQMTYTNVDLVLVGWQAPPGARSMTPCSGSSDLYLVTRHADV 1139
DB 1080 TVYHGAGSKTLIAGPKGPIQMTYTNVDLVLVGWQAPPGARSMTPCSGSSDLYLVTRHADV 1139
QY 1140 IPVRRRGRSGLLSRPPVSVYKSGSGGPLLCPSGHHVGVFRAAVCTRGVAKAVDFIPE 1199
DB 1140 IPVRRRGRSGLLSRPPVSVYKSGSGGPLLCPSGHVAVLFRFAAVCTRGVAKAVDFIPE 1199

QY 1200 SMETTWRSVFTDSTPPAVPOTFOVAHLHAPTGSKSTKVPAAYAAQGYKVLVLPNSVA 1259
DB 1200 SMETTWRSVFTDSTPPAVPOTFOVAHLHAPTGSKSTKVPAAYAAQGYKVLVLPNSVA 1259
QY 1260 ATLGFGAYNSKAHGDIPNRTGVRTITTCGSIITYSYGKFLADGGCGSGAYDIIICDECH 1319
DB 1260 ATLGFGAYNSKAHGDIPNRTGVRTITTCGPIITYSYGKFLADGGCGSGAYDIIICDECH 1319
QY 1320 STDSTTILGIVTLDQAEATAGARLVVLTATPPGSVTVPHNIEBTGLNNGEIPPYGKA 1379
DB 1320 STDSTTILGIVTLDQAEATAGARLVVLTATPPGSVTVPHNIEEVALSNTGIEIPYGKA 1379
QY 1380 IPIEAIKGRGHLIFCHSKKCCDELAAKITGLGINAVAYRGIDVSVIPIPGDVVVATDA 1439
DB 1380 IPIEVIKGRGHLIFCHSKKCCDELAAKITGLGINAVAYRGIDVSVIPIPGDVVVATDA 1439
QY 1440 LMTGFTGDFSDVDCNCTVTQTVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGSGIY 1499
DB 1440 LMTGFTGDFSDVDCNCTVTQTVDFSLDPTFTTETTTVPQDAVRSORRGRTGRGSGIY 1499
QY 1500 RFVTPGERPSGMPDSSVLCCEYDAGCAWVELTPAETSVRLRAYLNTFGLPVCODHLEFWE 1559
DB 1500 RFVTPGERPSGMPDSSVLCCEYDAGCAWVELTPAETSVRLRAYLNTFGLPVCODHLEFWE 1559
QY 1560 SVPTGLTHIDAHPLSOTKQAGDNFPYLVAQTVCARAQAPPSQDMWKCLIRLKPETH 1619
DB 1560 SVPTGLTHIDAHPLSOTKQAGDNFPYLAAYQATVCARAQAPPSQDMWKCLIRLKPETH 1619
QY 1620 GPTPLLYRLGAVONEVLTHTPIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLITGTS 1679
DB 1620 GPTPLLYRLGAVONEVLTHTPIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLITGTS 1679
QY 1680 VVTGVRILLSGKPAVPDREVLVQEFDMEBECASQLPYIEQGMQLAEQFKQKALGLLOTA 1739
DB 1680 VVTGVRILLSGKPAVPDREVLVQEFDMEBECASQLPYIEQGMQLAEQFKQKALGLLOTA 1739
QY 1740 TKQAEAAAPVESKWRALFTFAKHMWNFTISGTOYLAGLSTLPGNPAIASLMAFTASITS 1799
DB 1740 TKQAEAAAPVESKWRALFTFAKHMWNFTISGTOYLAGLSTLPGNPAIASLMAFTASITS 1799
QY 1800 PLTQNTLLNIIIGWVAAQLAPPSAASAFVAGIAGAAVGSIGLGLVLDIILAGYAGV 1859
DB 1800 PLTQNTLLNIIIGWVAAQLAPPSAASAFVAGIAGAAVGSIGLGLVLDIILAGYAGV 1859
QY 1860 AGALVAFKWSGEVPESTEDLVNLLPALISPCALVGVVCAAILRRRHVGPGEAGVQVMNRL 1919
DB 1860 AGALVAFKWSGEVPESTEDLVNLLPALISPCALVGVVCAAILRRRHVGPGEAGVQVMNRL 1919
QY 1920 IAFASRGNHVSPTHYVPESDAAARVTOILSLTITQLLKRHOWINEDCSTPCSGWLRD 1979
DB 1920 IAFASRGNHVSPTHYVPESDAAARVTOILSLTITQLLKRHOWINEDCSTPCSGWLRD 1979
QY 1980 VWDWICTVLDFKTWLOSULLPRLPVGFPLSCQRYGKVMRGDGMOTTCPCGAQIAGHV 2039
DB 1980 VWDWICTVLDFKTWLOSULLPRLPVGFPLSCQRYGKVMRGDGMOTTCPCGAQIAGHV 2039
QY 2040 KNGSMRITVGPRTCNTWHTGTPPINAYTGTCTPSPAPNYGRALVRVAEEVEVTRVGDV 2099
DB 2040 KNGSMRITVGPRTCNTWHTGTPPINAYTGTCTPSPAPNYGRALVRVAEEVEVTRVGDV 2099
QY 2100 HYVTGMTTDNVKCPQVPAPEFFTEVDGVRILHRYPACKPLLRBDVTFQVGLNLYLVGSQ 2159
DB 2100 HYVTGMTTDNVKCPQVPAPEFFTEVDGVRILHRYPACKPLLRBDVTFQVGLNLYLVGSQ 2159
QY 2160 LPCPEPDVTVLTSMITDPSHITATKRLRARGSPPSLASSASQLSAPSLKATCTTHH 2219
DB 2160 LPCPEPDVTVLTSMITDPSHITATKRLRARGSPPSLASSASQLSAPSLKATCTTHH 2219
QY 2220 DSPADLIEANLLWRQEMGNITRVESKENVVILDSPEPLHAEGDEREISVAAEILKRSR 2279
DB 2220 DSPADLIEANLLWRQEMGNITRVESKENVVILDSPEPLHAEGDEREISVAAEILKRSR 2279
QY 2280 KFPSALPTIWARPDYNPPLLESWKDPDYVPPVHGCPLPPTKAPPIPPPRKRRTVVLTESN 2339

DB 2280 KFPPALPTIWARPDYNPPLLESWKDPDYVPPVHGCPLPPTKAPPIPPPRKRRTVVLTEST 2339
QY 2340 VSSALAEALATKTFGSGSSAVDSGTATATLPDLASDDGDKGSDVESYSSMPPLEGEGDDP 2399
DB 2340 VSSALAEALATKTFGSGSSAVDSGTATATPPDOASDYGDKGSDVESYSSMPPLEGEGDDP 2399
QY 2400 LSDGWSVTSEASDVVCCSMSYTWGTGALITPCAAEESKLPINPLSNLSLLRHHNNVYAT 2459
DB 2400 LSDGWSVTSEASDVVCCSMSYTWGTGALITPCAAEESKLPINPLSNLSLLRHHNNVYAT 2459
QY 2460 TSRSASLRKQKVTFRDLQVLDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSK 2519
DB 2460 TSRSASLRKQKVTFRDLQVLDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSK 2519
QY 2520 FGYGKADVRLSSRAVNHIRSWEDELLETETPITTTIMAKSEVFCVQEKGRKPARLI 2579
DB 2520 FGYGKADVRLSSRAVNHIRSWEDELLETETPITTTIMAKSEVFCVQEKGRKPARLI 2579
QY 2580 VFPDLGVRCEKVALYDVVSTLPQAVMGSSYGFQYSPKORVEFLVNTWKS KCPMGFSYD 2639
DB 2580 VFPDLGVRCEKVALYDVVSTLPQAVMGSSYGFQYSPKORVEFLVNTWKS KCPMGFSYD 2639
QY 2640 TRCFDSTVTESDIRVEESYIQCCLAPPEARQAIRSLTERLYIGGPLTNSKGQCYRRRCR 2699
DB 2640 TRCFDSTVTENDIRVEESYIQCCLAPPEARQAIRSLTERLYIGGPLTNSKGQCYRRRCR 2699
QY 2700 ASGVLTS CGNTLT CYLKATAACRAAKLQDCTMLVNGDDLWVICESAGTOEDAAALRAFT 2759
DB 2700 ASGVLTS CGNTLT CYLKATAACRAAKLQDCTMLVNGDDLWVICESAGTOEDAAALRVFT 2759
QY 2760 EAMTRISAPGPPDPOPEYDLELITSCSSNVSVAHADAGKRVVYLTDRDPTTPLARAAWETA 2819
DB 2760 EAMTRISAPGPPDPOPEYDLELITSCSSNVSVAYDASGRKRVYLTDRDPTTPLARAAWETA 2819
QY 2820 RHTPINSWLGNIIMYAPTLWARMLMTHFILLSAQLEKALDCCQIYGACYSIEPLDLP 2879
DB 2820 RHTPINSWLGNIIMYAPTLWARMLMTHFILLSAQLEKALDCCQIYGACYSIEPLDLP 2879
QY 2880 QIIRLHGLSAFTLHSPCEINRVASCLRKGLVPLRTWRHRARSVRAKLLSQGGRRAAT 2939
DB 2880 QIIRLHGLSAFTLHSPCEINRVASCLRKGLVPLRVWRHRARSVRAKLLSRGGRRAAT 2939
QY 2940 CGRYLFNVAVRTKLTPTIPAASQDLSCGFVAGYGGDIYHLSRARPRWFLCULLLS 2999
DB 2940 CGRYLFNVAVRTKLTPTIPAASQDLSCGFVAGYGGDIYHLSRARPRWFLCULLLS 2999
QY 3000 VGVGIYLLPNR 3010
DB 3000 VGVGIYLLPNR 3010

RESULT 13
Q9QIX6_9HEPC PRELIMINARY; PRT; 3010 AA.
ID Q9QIX6_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIX6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD8-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
activity."; Virology 263:244-253(1999).

RN [2]
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish P., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF165059; AAD56194.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; P00246; P00246.
DR PIR; P00254; P00254.
DR PIR; P00804; P00804.
DR PIR; P50329; P50329.
DR HSP; Q8JYS1; 1CWX.
DR SMR; Q9QIX6; 1029-1657, 2008-2170, 2420-2949.
DR MEROPS; S29.002; -; C: viral capsid; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003168; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0005198; P: structural molecule activity; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_ U39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327471 MW; 4613744EC4D4A013 CRC64;
Query Match 95.9%; Score 15356; DB 2; Length 3010;
Best Local similarity 95.0%; Pred. No. 0;
Matches 2861; Conservative 69; Mismatches 80; Indels 0; Gaps 0;
QY 1 MSTNPKPQRTKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATRKASRSQPRG 60

Db 1 MSTNPKPQRTKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSRSQPRG 60
QY 61 RRQIPKARPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDPRRRSNL 120
Db 61 RRQIPKARPEGRWAQPGYPWPLYGNEGWAGWLLSPRGSRPSWGPTDPRRRSNL 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEGVNATCNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDSVNYATCNLPGCSFSIFLLA 180
QY 181 LLSCLTTPASAYEVNRVNGIYHVNTDCSNSIYVEAADVIHMTPCVPCVQVQBGSSRCW 240
Db 181 LLSCLTTPASAYEVNRVNGIYHVNTDCSNSIYVEAADIIHMTPCMPCVREDNISRCW 240
QY 241 ALPTLAARNASVPTTIRRHVDLLVGTAAPCSAMYVDLCGSIFLYSQSLTFSPRHET 300
Db 241 ALPTLAARNISVPTTIRRHVDLLVGTAAPCSAMYVDLCGSLVLSQSLTFSPRHET 300
QY 301 VQDCNCSIYPGHVSGHRMAMDMNMWSPPTALVVSQLLRIPQAVVDMVAGAHWGLAGLA 360
Db 301 VQDCNCSIYPGHVSGHRMAMDMNMWSPPTAALVVSQLLRIPQAVVDMVAGAHWGLAGLA 360
QY 361 YYSVMGNWAKVLIIVALLFAGVDGETHTRVAGHTTSGFTSLPSSGASOKIQLVNTNGSW 420
Db 361 YYSVMGNWAKVLIIVALLFAGVDGTHVTGQTQARSAILTSFSLTVGPSQKQLVNTNGSW 420
QY 421 HINRTALNCNDSLOTGFFAALFYAHKNSGCCPERMASCRPIDMFAOGWGPIITTKNSS 480
Db 421 HINRTALNCNDSLOTGFLAALFYTHRFNAGSCPERMASCRPIDTDOGWGPIITTAEPNP 480
QY 481 DQPCYCHYAPRPGCVVPASQVCPVCFPTSPVVVGTDRSGVPTYSWGENETDMLLN 540
Db 481 DQPCYCHYAPRPGCIEFPASKVCGPVCFPTSPVVVGTDRFGVPTYNWGENETDMLLN 540
QY 541 NTRPPQGNWFGCTWNSTGFTKTCGGPPCNIGVGNRTLICPTDCFRKHPKATYTKCSG 600
Db 541 NTRPPQGNWFGCTWNSTGFTKTCGGPPCNIGAGNNTLICPTDCFRKHPKATYTKCSG 600
QY 601 PWTLPCLVDYPIELWHYPCPLNESIFKVRMYGVGVEHRLNAA CNWTRGECNLEDRDRS 660
Db 601 PWTLPCLVDYPIELWHYPCPLNFTIFKVRMYGVGVEHRLSAA CNWTRGECNLEDRDRS 660
QY 661 ELSPLLLSTTEWQILPCAFITLPAALSTGLIHLHQNIVDVQVLYGVGSFAVFAIKWEYIL 720
Db 661 ELSPLLLSTTEWQILPCSFITLPAALSTGLIHLHQNIVDVQVLYGVGSFAVFAIRWEYVL 720
QY 721 LFLLLADARVCACIMWMLLIAQAEALLENIVLNAASVAGAHGILSLFLVFFCAAWYIKG 780
Db 721 LFLLLADARVCACIMWMLLIAQAEALLENIVLNAASVAGAHGILSLFLVFFCAAWYIQG 780
QY 781 RLAPGAAYAFYGVWPLILLILALPRAVALDRENAASCGGAVLVCLVFLTLPYVYKVELT 840
Db 781 RLVPGAAYALYGVWPLILLILALPRAVALDRENAASCGGAVFVGLAFELTLPYVYKVELA 840
QY 841 RLIIWLVQYFTRAEAHQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILGLPLM 900
Db 841 RLIIWLVQYFTRAEAHQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLAILFGPLM 900
QY 901 VLOAGITRVPYFVRAQGLIRACMLVRKVAGHYVQVFMFKLGALTGTVYVYVNHLPDRWA 960
Db 901 VLOAGITRVPYFVRAHGLIRACMLVRKVAGHYVQVFMFKLAALTGTVYVYVHLPDQWA 960
QY 961 HAGLRDLAVAVEPVVFSMETKVIITWGDATACGDIILGLPVSARRGKEIFLGPADSLEG 1020
Db 961 HSLGRDLAVAVEPVVFSMETKVIITWGDATACGDIILGLPVSARRGKEIILGPADSLEG 1020
QY 1021 QGWRLLAPITAYSQOTRGVLCIITSLTGRDNQVEGEVQVVSSTATQSFATCINGVCWT 1080
Db 1021 QGWRLLAPITAYSQOTRGVLCIITSLTGRDNQVEGEVQVVSSTATQSFATCINGVCWT 1080
QY 1081 VYHAGSKTLAPKGPITQMTYTNVDLDLVGWOAPPGARSMTPTCSGSSDLVLYVTHADVI 1140

Db 1081 VFHAGSKTLAGKGPITOMYTNVDLGVQAPPGARSMTPTCTGSSDLYLVRHADVI 1140
Qy 1141 PVRRGDSRGSLLSPRPVSYLKGSSGGLLCPGSHVGVFRAAVCTRGVAKAVDFIPVES 1200
Db 1141 PVRRGDSRGSLLSPRPVSYLKGSSGGLLCPGSHVGVFRAAVCTRGVAKAVDFIPVES 1200
Qy 1201 METTMRSPVFTDNSTPAPVQTFQVAHLHAPTGSKSTKVPAAAYAOQYKVLVLPNSVAA 1260
Db 1201 METTMRSPVFTDNSTPAPVQTFQVAHLHAPTGSKSTKVPAAAYAOQYKVLVLPNSVAA 1260
Qy 1261 TLGFGAYMSKANGIDNIKTGRTITGSGITTYSTYCKELADGGCGGAYDIIICDECHS 1320
Db 1261 TLGFGAYMSKANGIDNIKTGRTITGSGITTYSTYCKELADGGCGGAYDIIICDECHS 1320
Qy 1321 TOSTTILGIGTVLDOAETAGARLVLAATATPGSVTPHPNTEIEIGLSNNGEIPFYGKAI 1380
Db 1321 TOSTTILGIGTVLDOAETAGARLVLAATATPGSVTPHPNTEIEIGLSNNGEIPFYGKAI 1380
Qy 1381 PTEAIKGGRHILFCHSKKKKDELAALKLIGLGNVAYYRGDVSVPPIGDVVVVAATDAL 1440
Db 1381 PIEVIKGGRHILFCHSKKKKDELAALKLIGLGNVAYYRGDVSVPPIGDVVVVAATDAL 1440
Qy 1441 MTGFTGDFSDVDCNTCVTQVDFSLDPTTETTTVPQDAVRSORRGTRGRSGIYR 1500
Db 1441 MTGFTGDFSDVDCNTCVTQVDFSLDPTTETTTVPQDAVRSORRGTRGRSGIYR 1500
Qy 1501 FVTPGRRPSGMFSDVCECYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWES 1560
Db 1501 FVTPGRRPSGMFSDVCECYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWES 1560
Qy 1561 VFTGLTHIDAHFLOSOTKQAGNDPPYLVAOYTCARAQAPPPSDQWQKCLILKPLTHG 1620
Db 1561 VFTGLTHIDAHFLOSOTKQAGNDPPYLVAOYTCARAQAPPPSDQWQKCLILKPLTHG 1620
Qy 1621 PTPLLYRLGAVQNEVILTHPIKYIMACMSADLEVTSTVWLVGGVLAALAAAYCLITGTV 1680
Db 1621 PTPLLYRLGAVQNEVILTHPIKYIMACMSADLEVTSTVWLVGGVLAALAAAYCLITGTV 1680
Qy 1681 VIVGRILSGKPAVPDPREVLYGEFDEMEECASOLPYIEQGMQLAEQFKQKALGLQAT 1740
Db 1681 VIVGRILSGKPAVPDPREVLYGEFDEMEECASOLPYIEQGMQLAEQFKQKALGLQAT 1740
Qy 1741 KQAEAAAPVVEKRWALLETFAWKHMNFIISGQIYLAGLSTLPGNPAIASLMAFTASITSP 1800
Db 1741 KQAEAAAPVVEKRWALLETFAWKHMNFIISGQIYLAGLSTLPGNPAIASLMAFTASITSP 1800
Qy 1801 LTTQNTLLFNILGSGWAAQAPSAASAFVAGTAGAAGVSGIGLKVLDVILAGYGAGVA 1860
Db 1801 LTTQNTLLFNILGSGWAAQAPSAASAFVAGTAGAAGVSGIGLKVLDVILAGYGAGVA 1860
Qy 1861 GALVAPKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLI 1920
Db 1861 GALVAPKVMGSEVPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLI 1920
Qy 1921 AFASRGNHVSPTHVVPESDAARVTQILSSITITQLLKRHLQWINEPCSTPCSGSWLRDV 1980
Db 1921 AFASRGNHVSPTHVVPESDAARVTQILSSITITQLLKRHLQWINEPCSTPCSGSWLRDV 1980
Qy 1981 WDWICTVLTPKTLQSKLPLRPGVPLSCQRYKGVWRGDGIMQITPCGAQIAGHVK 2040
Db 1981 WDWICTVLSDPKTLQSKIMPLRPGVPLSCQRYKGVWRGDGIMQITPCGAQIAGHVK 2040
Qy 2041 NGSMRIIVGPRCTSNWHTGTFPINATYTCPTSPAPNYSRALWVAEEVYVTRVGDHF 2100
Db 2041 NGSMRIIVGPRCTSNWHTGTFPINATYTCPTSPAPNYSRALWVAEEVYVTRVGDHF 2100
Qy 2101 YVTGWTNDNVKPCQVPAPEFTEVDGVRHLRYAPACKPLREDVTFQVGLNQYLVGSQ 2160
Db 2101 YVTGWTNDNVKPCQVPAPEFTEVDGVRHLRYAPACKPLREDVTFQVGLNQYLVGSQ 2160
Qy 2161 PCEPEPDVTLTSMLTDPSSHITAEAKRRLARGSPPSLASSASQLSAPSKATCTTHHD 2220
Db 2161 PCEPEPDVTLTSMLTDPSSHITAEAKRRLARGSPPSLASSASQLSAPSKATCTTHHD 2220

Qy 2221 SPADLIIEANLWROEMGNIITRVESENKVVILDSFEPHHAEGDEREISVAAEILRKSPK 2280
Db 2221 SPADLIIEANLWROEMGNIITRVESENKVVILDSFEPHHAEGDEREISVAAEILRKSPK 2280
Qy 2281 FPSALPIWARPDYNNPPLLESWKDPDVPVPPVHGCPPLPTTKAPPIPPPRKRRTVVLTESNV 2340
Db 2281 FPSALPIWARPDYNNPPLLESWKDPDVPVPPVHGCPPLPTTKAPPIPPPRKRRTVVLTESIV 2340
Qy 2341 SSALAEALATKTFGSSGSSAVDSGTATAPDLASDDGDKGSDVESYSSMPPLEGEPPDL 2400
Db 2341 SSALAEALATKTFGSSGSSAVDSGTATAPDLASDDGDKGSDVESYSSMPPLEGEPPDL 2400
Qy 2401 SDGSMSTVSEASESDVVCMSMTWTGALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
Db 2401 SDGSMSTVSEASESDVVCMSMTWTGALITPCAAEESKLPINPLNSLLRHHNMVYATT 2460
Qy 2461 SRSASLRQKVTFFDLRLQVLDHRYDVLKEMKAKASTVKAKLSIEEACKLTPHSAKSKF 2520
Db 2461 SRSASLRQKVTFFDLRLQVLDHRYDVLKEMKAKASTVKAKLSIEEACKLTPHSAKSKF 2520
Qy 2521 GYGAKDVRNLSSRAVNHRSVWEDLLEDTEPTDITIMAKSEVFCVQPEKGRKPARLIV 2580
Db 2521 GYGAKDVRNLSSRAVNHRSVWEDLLEDTEPTDITIMAKSEVFCVQPEKGRKPARLIV 2580
Qy 2581 FPDGLFVRCVKEMALYDVVSTLTPQAVMGSSYGFQYSPKQRFVFLVITWKKKCPMGFSYDT 2640
Db 2581 FPDGLFVRCVKEMALYDVVSTLTPQAVMGSSYGFQYSPKQRFVFLVITWKKKCPMGFSYDT 2640
Qy 2641 RCFDSTVTESDIRVESIYQCCDLAPEARQAIISLTERLYIGGPLTNSKGQNGCYRRCRA 2700
Db 2641 RCFDSTVTESDIRVESIYQCCDLAPEARQAIISLTERLYIGGPLTNSKGQNGCYRRCRA 2700
Qy 2701 SGVLTTSCGNTLCYLKATAACRAAKLOCTMLVNGDDLVVICESAGTQSDAALRAFTE 2760
Db 2701 SGVLTTSCGNTLCYLKATAACRAAKLOCTMLVNGDDLVVICESAGTQSDAALRAFTE 2760
Qy 2761 AMTRYAPCPDPQPPQDYDELELITSCSSNVSAHDASKRVVYILTRDPTTTPALAAAWETAR 2820
Db 2761 AMTRYAPCPDPQPPQDYDELELITSCSSNVSAHDASKRVVYILTRDPTTTPALAAAWETAR 2820
Qy 2821 HTPINSWLGNIIMYAPTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQ 2880
Db 2821 HTPINSWLGNIIMYAPTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQ 2880
Qy 2881 IIERLHGLSAFTLHYSYSPGEINRVAASCLRLKGLVPPPLTRWHRARSVRKALLSQGGAATC 2940
Db 2881 IIERLHGLSAFTLHYSYSPGEINRVAASCLRLKGLVPPPLTRWHRARSVRKALLSQGGAATC 2940
Qy 2941 GRVLFNVAVRTKLTLPFAASQIDLSGWVAGYSGGDIYHSLSRARPFRFPCLCLLLSV 3000
Db 2941 GRVLFNVAVRTKLTLPFAASQIDLSGWVAGYSGGDIYHSLSRARPFRFPCLCLLLSV 3000
Qy 3001 GVGIYLLPNR 3010
Db 3001 GVGIYLLPNR 3010

RESULT 14
Q9J3H9_9HEPC PRELIMINARY; PRT: 3010 AA.
ID Q9J3H9_9HEPC
AC Q9J3H9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD13;


```
QY 1261 TLGFGAYMKAHGIDNIRITGVRTITTTGSIYTVSTYVKFLADGGCGGGAYDIIICDECHS 1320
DB 1261 TLGFGAYMKAHGVDNIRITGVRTITTTGAPITTYSTYVKFLADGGCGGGAYDIIICDECHS 1320
QY 1321 TDSSTTILGTVLDAQETAGARLWVLATATPPGSGVTPHPNIEEIGLSNNGEIPFYGKAI 1380
DB 1321 TDSSTTILGTVLDAQETAGARLWVLATATPPGSGVTPHPNIEEALSNTGEIPFYGKAI 1380
QY 1381 PIBAIGKRHLIFCHSKKKDELAALKLGLGLNAVAYYRGLDVSVPIPGDVVVVATDAL 1440
DB 1381 PIETIKGRHLIFCHSKKKDELAALKLGLGLNAVAYYRGLDVSVPIPGDVVVVATDAL 1440
QY 1441 MTGFTGDFSDVINDCNTVTVDFSLDPTFTIETTTVPDASRSORRGRTGRGSIYR 1500
DB 1441 MTGYTGDFSDVINDCNTVTVDFSLDPTFTIETTTVPDASRSORRGRTGRGSIYR 1500
QY 1501 FVTPGERPSGMFSSVLCYDAGCAYELTPAETSURLRAYLNTPLGLPVCQDHLFWES 1560
DB 1501 FVTPGERPSGMFSSVLCYDAGCAYELTPAETSURLRAYLNTPLGLPVCQDHLFWES 1560
QY 1561 VFTGLTHIDAHFLSQTQAGDNFPYLVAQATVCARAQAPPPSWDQWKCILRLKPTLHG 1620
DB 1561 VFTGLTHIDAHFLSQTQAGDNFPYLVAQATVCARAQAPPPSWDQWKCILRLKPTLHG 1620
QY 1621 PTPLLYRLGAVONEVILTHPIKYIMACMSADLEVTVTWLVGGVLAALAAAYCLITGVS 1680
DB 1621 PTPLLYRLGAVONEVILTHPIKYIMACMSADLEVTVTWLVGGVLAALAAAYCLITGVS 1680
QY 1681 VIIGRIILSKPAVPDPREVLYOEFDEMERCASOLPYIEOGMOLAEQFKKALGLLQAT 1740
DB 1681 VIIGRIILSKPAVPDPREVLYOEFDEMERCASOLPYIEOGMOLAEQFKKALGLLQAT 1740
QY 1741 KQAEAAAPVYESKRALETFWAKHMNFISGIOYLAGLSTPLGNPAIASMAFTASTSP 1800
DB 1741 KQAEAAAPVYESKRALETFWAKHMNFISGIOYLAGLSTPLGNPAIASMAFTASTSP 1800
QY 1801 LTTONTLLFNILGWWAAQAPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVA 1860
DB 1801 LTTONTLLFNILGWWAAQAPSAASAFVAGIAGAAVGSIGLGKVLVDILAGYGAGVA 1860
QY 1861 GALVAFKMSGEVPSSTEDLVNLLPAILSPGALVGVVCAILLRHVGPGEVAVOMNRLI 1920
DB 1861 GALVAFKMSGEVPSSTEDLVNLLPAILSPGALVGVVCAILLRHVGPGEVAVOMNRLI 1920
QY 1921 AFASRGNHVSPTHVVPESDAAARVTQILSSITITQLLRLHQWINEPCSTPCSGSWLRDV 1980
DB 1921 AFASRGNHVSPTHVVPESDAAARVTQILSSITITQLLRLHQWINEPCSTPCSGSWLRDV 1980
QY 1981 WDWICTVLDTFKWLQSKLLPRLPGVDFLSCQRYKGVWRGDMQITTCPCGAQIAGHVK 2040
DB 1981 WDWICTVLDTFKWLQSKLLPRLPGVDFLSCQRYKGVWRGDMQITTCPCGAQIAGHVK 2040
QY 2041 NGSNRIIVGPRTCSTNWTGTPINAYTTGCTTPSPAPNYSRALWRAVEEYVETRVGDEH 2100
DB 2041 NGSNRIIVGPRTCSTNWTGTPINAYTTGCTTPSPAPNYSRALWRAVEEYVETRVGDEH 2100
QY 2101 YVTGMTTNDVNCQVPAPEPFTVDGVRLHRYAPACKPLLRREDVTVQVGLNOYLVSQ 2160
DB 2101 YVTGMTTNDVNCQVPAPEPFTVDGVRLHRYAPACKPLLRREDVTVQVGLNOYLVSQ 2160
QY 2161 PCEPEPDVTLTSMLTDPFSHITAEAKRRLARGSPPLSASSASQLSAPSKATCTTHHD 2220
DB 2161 PCEPEPDVTLTSMLTDPFSHITAEAKRRLARGSPPLSASSASQLSAPSKATCTTHHD 2220
QY 2221 SPDADLLEANLLWQEMGNNITRVESENKVVILDSFPLHAEGEREISVAEILRSRK 2280
DB 2221 SPDADLLEANLLWQEMGNNITRVESENKVVILDSFPLHAEGEREISVAEILRSRK 2280
QY 2281 FPSALPTIWARDYNNPILLESKWDYDYPVPHGCLPPTKAPPIPPPPRKRKTIVLTSSNV 2340
DB 2281 FPSALPTIWARDYNNPILLESKWDYDYPVPHGCLPPTKAPPIPPPPRKRKTIVLTSSNV 2340
QY 2341 SSALAEALATKTFGSSGSSAVDSGTATALPDLASDDGDKGSDVESYSSMPPLEGEFGPDL 2400
```

```
DB 2341 SAALAEALATKTFGSSSSAADSATATAPDQASDDGKESDVESYSSMPPLEGEFGPDL 2400
QY 2401 SDGSMSTVSEEAESDVVCCSMSTYTTGALITPCAABESKLPIINPLSNLLRHNMVYATT 2460
DB 2401 SDGSMSTVSEEAESDVVCCSMSTYTTGALITPCAABESKLPIINPLSNLLRHNMVYATT 2460
QY 2461 SRSASLRQKKVTFDRLOVLDDHYRDVLEKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
DB 2461 SRSAGURQKKVTFDRLOVLDDHYRDVLEKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKF 2520
QY 2521 GYGAKDVRNLSRAVNHRSVMEDELTETPTDITTIMAKSEYFCVQPEKGRKPARLIY 2580
DB 2521 GYGAKDVRNLSRAVNHRSVMEDELTETPTDITTIMAKSEYFCVQPEKGRKPARLIY 2580
QY 2581 FPDGLVVRCEKMALYDVVSTLPOAVMGSSYGFQSPQRVFEFLVNTWKSXKCPMGFSYDT 2640
DB 2581 FPDGLVVRCEKMALYDVVSTLPOAVMGSSYGFQSPQRVFEFLVNTWKSXKCPMGFSYDT 2640
QY 2641 RCFDSTVTTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLTNSKGONCYRRCRA 2700
DB 2641 RCFDSTVTTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLTNSKGONCYRRCRA 2700
QY 2701 SGVLTTSCGNLTTCYLKATAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALAFTE 2760
DB 2701 SGVLTTSCGNLTTCYLKATAACRAAKLQDCTMLVNGDDLVIICESAGTQEDAAALAFTE 2760
QY 2761 AMTRYSAAPPDGPPEYDLELITSCSSNSVSAHDAGSKRVYVLTTRDPTTPPLARAAMETAR 2820
DB 2761 AMTRYSAAPPDGPPEYDLELITSCSSNSVSAHDAGSKRVYVLTTRDPTTPPLARAAMETAR 2820
QY 2821 HTPIINSWLGNIINVAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQ 2880
DB 2821 HTPIINSWLGNIINVAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQ 2880
QY 2881 IIERLHGLSAFTLHSPGSEINRVASCLRKLGVPPPLRTWRHRARSVRKLLSQSGRAATC 2940
DB 2881 IIERLHGLSAFTLHSPGSEINRVASCLRKLGVPPPLRTWRHRARSVRKLLSQSGRAATC 2940
QY 2941 GRYLFNWAVRTKLTPIPAASQLDLSGWFVAGYSGGDIYHSISRARPFRPPCLLLLSV 3000
DB 2941 GRYLFNWAVRTKLTPIPAASQLDLSGWFVAGYSGGDIYHSISRARPFRPPCLLLLSV 3000
QY 3001 GVGIYLLPNR 3010
DB 3001 GVGIYLLPNR 3010
```

RESULT 15

```
Q9DTE6_9HEPC PRELIMINARY; PRT; 3010 AA.
ID Q9DTE6_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9DTE6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohca Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RL Hepatol. Res. 20:161-171(2001).
DR EMBL; AB049091; BAB18804.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSSP; Q81755; 1DXP.
```

DR SMR; Q9DTB6; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005534; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR02522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_039_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327043 MW; 3807DC6879684C95 CRC64;

Query Match 95.8%; Score 15330; DB 2; Length 3010;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 2850; Conservative 75; Mismatches 85; Indels 0; Gaps 0;

QY 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVLLPRGPRLGVRATRKASERSQPRG 60
DB |||||
QY 1 MSTNPKPQKTKRNTNRRPDQVKFPGGGQIVGGVLLPRGPRLGVRATRKTSERSQPRG 60
DB |||||

QY 61 RRQPIPKARRPEGRAWAQPGYWPPLYGNEGLGWAGWLLSPGRSRPSWGPDPDRRSRNLG 120
DB |||||

QY 61 RRQPIPKARQPEGRAWAQPGYWPPLYGNEGLGWAGWLLSPGRSRPSWGPDPDRRSRNLG 120
DB |||||

QY 121 KVIDTLTCGFADLMGVPIPLVGAPLGAARALAHGVRLVDGVNYATGNLPGCSFSIFLIA 180
DB |||||

QY 121 KVIDTLTCGFADLMGVPIPLVGAPLGAARALAHGVRLVDGVNYATGNLPGCSFSIFLIA 180
DB |||||

QY 181 LLSCLTIPASAYEVNRVSGIYHVTDNCSNIIYEAADVIMHTPGCVPCVQEGNSRCWV 240
DB |||||

QY 181 LLSCLTIPASAYEVNRVSGIYHVTDNCSNIIYEAADVIMHAAPGCVPCVRENNSRCWV 240
DB |||||

QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVGDLGCSIFLVSQLFPSPRRHET 300
DB |||||

QY 241 ALTPTLAARNASVPTTIRRHVDLLVGTAAFCSAMVVGDLGCSVFLISQLFPSPRRHET 300
DB |||||

QY 301 VQDCNCSLYPGHVSGRHAWDMNMWSPPTALVVSQLLRIPQAVDMVAGAHGVLGAGLA 360
DB |||||

QY 301 VQDCNCSLYPGHVTGRHAWDMNMWSPPTAALVVSQLLRIPQAVDMVAGAHGVLGAGLA 360
DB |||||

QY 361 YYSVMGNWAKVLIVALLFAGVDGTHVTGGAQGRSAFRLTSFFTSAGSQIKLINTNGSW 420
DB |||||

361 YYSVMGNWAKVLIVALLFAGVDGTHVTGGAQGRSAFRLTSFFTSAGSQIKLINTNGSW 420
DB |||||

QY 421 HINRTALNCNDSLOTGPFALFVAHKENSGCERPMASCRPIDWFAQGWGPIITTKPNS 480
DB |||||

QY 421 HINRTALNCNDSLOTGPFALFVAHKENSGCERPMASCRPIDWFAQGWGPIITTKPNS 480
DB |||||

QY 481 DQRYCHYAPRPCGVVPASQVGPVCFPTSPVVVGTTRDSRGSVPTYSWGENETDMLLN 540
DB |||||

QY 481 DQRYCHYAPRPCGVVPASQVGPVCFPTSPVVVGTTRDSRGSVPTYSWGENETDMLLN 540
DB |||||

QY 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGVGNRTLICPTDCFRKHPEATYTKCGSG 600
DB |||||

QY 541 NTRPPQGNWFGCTWMNSTGFTKTCGGPPCNIGVGNRTLICPTDCFRKHPEATYTKCGSG 600
DB |||||

QY 601 PWTLPCLVDYPRYELMHPCTLNFSIFKVRMYGVGVEHRLNAACNWTGERCNLEDRDRS 660
DB |||||

QY 601 PWTLPCLVDYPRYELMHPCTLNFSIFKVRMYGVGVEHRLNAACNWTGERCNLEDRDRS 660
DB |||||

QY 661 ELSPLLLSTTEWQILPCAFITLPAALSTGLIHLHONIVDVQVLYGVGSFAVFAIKWEYIL 720
DB |||||

QY 661 ELSPLLLSTTEWQILPCAFITLPAALSTGLIHLHONIVDVQVLYGVGSFAVFAIKWEYIL 720
DB |||||

QY 721 LLFLLADARVCACACWMLLIAQAAALNLVILNAASVAGAHGILSLFVFCFAAWYIKG 780
DB |||||

QY 721 LLFLLADARVCACACWMLLIAQAAALNLVILNAASVAGAHGILSLFVFCFAAWYIKG 780
DB |||||

QY 781 RLAFGAAYAFYGVWPLLLLLLALPPRAYALDREMAASCGGAVLYGLVPLTSLSPYKVPILT 840
DB |||||

QY 781 RLAFGAAYAFYGVWPLLLLLLALPPRAYALDREMAASCGGAVLYGLVPLTSLSPYKVPILT 840
DB |||||

QY 841 RLIWLVQVITRAEAHQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLALITGLPM 900
DB |||||

QY 841 RLIWLVQVITRAEAHQVWVPLNVRGGRDAIILLTCAVHPELIFDITKLLALITGLPM 900
DB |||||

QY 901 VLOAGITRVPYFVRAOGLIRACMLVRKVAGHYVQVPMKLGALITGVYVNHLPRLDWA 960
DB |||||

QY 901 VLOAGITRVPYFVRAOGLIRACMLVRKVAGHYVQVPMKLGALITGVYVNHLPRLDWA 960
DB |||||

QY 961 HAGLRDLAVAVEPVVFSAMETKVITWGDATACGDIILGLPVSARRGKEIFLGPADSLEG 1020
DB |||||

QY 961 HAGLRDLAVAVEPVVFSAMETKVITWGDATACGDIILGLPVSARRGKEIFLGPADSLEG 1020
DB |||||

QY 1021 OGWRLLAPITAYSOOTRGVLGCIITSLTRDKNVEGEVQVSTATQSFATCINGVCWT 1080
DB |||||

QY 1021 OGWRLLAPITAYSOOTRGVLGCIITSLTRDKNVEGEVQVSTATQSFATCINGVCWT 1080
DB |||||

QY 1081 VYHAGSKTLAGPKGPIQMTYTNVDLVLGVQAPPGARSMTPCSCGSSDLYLVTTRHADVI 1140
DB |||||

QY 1081 VYHAGSKTLAGPKGPIQMTYTNVDLVLGVQAPPGARSMTPCSCGSSDLYLVTTRHADVI 1140
DB |||||

QY 1141 PVRRGDSRGSLLSPRVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200
DB |||||

QY 1141 PVRRGDSRGSLLSPRVSYLKGSSGGPLLCPSGHVGVFRAAVCTRGVAKAVDFIPVES 1200
DB |||||

QY 1201 METTMRSPVFTDNTSPPAVPOTFOVAHLHAPGSGSKTKVPAAYAAQGVKVLVLPNSVAA 1260
DB |||||

QY 1201 METTMRSPVFTDNTSPPAVPOTFOVAHLHAPGSGSKTKVPAAYAAQGVKVLVLPNSVAA 1260
DB |||||

QY 1261 TLGFCAYMSKAGHDIPNIRTGVRTITTCGSIITYSGKFLADGGCGGGAYDIIICDECHS 1320
DB |||||

QY 1261 TLGFCAYMSKAGHDIPNIRTGVRTITTCGSIITYSGKFLADGGCGGGAYDIIICDECHS 1320
DB |||||

QY 1321 TDSSTILGIGTVDQAEATAGARLVVATATPPGSVTVPHNIEETGLSNNGIPIFYGKAI 1380
DB |||||

QY 1321 TDSSTILGIGTVDQAEATAGARLVVATATPPGSVTVPHNIEETGLSNNGIPIFYGKAI 1380
DB |||||

QY 1381 PTEATKGRHILFCHSKKKCDLAAKLTGLGNVAVYVGLDVSVIPIGDDVVVATDAL 1440
DB |||||

QY 1381 PTEATKGRHILFCHSKKKCDLAAKLTGLGNVAVYVGLDVSVIPIGDDVVVATDAL 1440
DB |||||

QY 1441 MTGFTGDFSDSDVCNCTVTQTVDFTFTTFTTTPQDAVRSQRGRGTSYR 1500
DB |||||

Db 1441 MTGYTGDFSVI DNTCVTQTVDFSLDPTFTIETTTVPQDAVSRTQBRGRTGRGRGIYR 1500

QY 1501 FVTPGERPSGMFSSVLCYDCYAGCWAYELTPTAETS VRLRAYLNTPLGPVQCQDHLFEWES 1560

Db 1501 FVTPGERPSGMFSSVLCYDCYAGCWAYELTPTAETS VRLRAYLNTPLGPVQCQDHLFEWES 1560

QY 1561 VFTGLTHIDAHFISQTKQAGDNFPYLVAQYATVCARAQAAPPSWDQMWKCLIRLKP TLHG 1620

Db 1561 VFTGLTHIDAHFISQTKQAGDNFPYLVAQYATVCARAQAAPPSWDQMWKCLIRLKP TLHG 1620

QY 1621 PTLPLVRLGAVQNEVILTHPIKYI MACMSADLEVTSTWVLVGGVLAALAAACLTITGSV 1680

Db 1621 PTLPLVRLGAVQNEVILTHPIKYI MACMSADLEVTSTWVLVGGVLAALAAACLTITGSV 1680

QY 1681 VIYGRITILSGKPAVVPDREVLVYOEFDMEBECASQLPYIEQGMQLAEQFKOKALGLLQTAT 1740

Db 1681 VIYGRITILSGKPAVVPDREVLVYOEFDMEBECASQLPYIEQGMQLAEQFKOKALGLLQTAT 1740

QY 1741 KOAEAAAPVVESKWRALLETFWAKHMNFI SGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800

Db 1741 KOAEAAAPVVESKWRALLETFWAKHMNFI SGIQYLAGLSTLPGNPAIASLMAFTASITSP 1800

QY 1801 LTTQNTLLFNILGWAQAOLAPSAASAFVGCAGTAGAAGVSGIGLKVLDVILAGYAGGVA 1860

Db 1801 LTTQNTLLFNILGWAQAOLAPSAASAFVGCAGTAGAAGVSGIGLKVLDVILAGYAGGVA 1860

QY 1861 GALVAFKVMGSEVPSTEDLVNLLPAIILSPGALVGVVCAAILRHVGPGECAVQMMNRLI 1920

Db 1861 GALVAFKVMGSEVPSTEDLVNLLPAIILSPGALVGVVCAAILRHVGPGECAVQMMNRLI 1920

QY 1921 AFASRGNHVSPTHVVPESDAAARVTQILSSITITOLLKRLHOWINEDCSTPCSGSWLRDV 1980

Db 1921 AFASRGNHVSPTHVVPESDAAARVTQILSSITITOLLKRLHOWINEDCSTPCSGSWLRDV 1980

QY 1981 WDNICTVLTDKFWLQSKLPLRGVDFLSCORGYKGVWRGDGIMQITTCPCGAOIAGHVK 2040

Db 1981 WDNICTVLTDKFWLQSKLPLRGVDFLSCORGYKGVWRGDGIMQITTCPCGAOIAGHVK 2040

QY 2041 NGSMRI VGPRTCSNTWGTFFPINAYTTGCTPPSPAPNYSRALMRVAAEYVEVTRVGDH 2100

Db 2041 NGSMRI VGPRTCSNTWGTFFPINAYTTGCTPPSPAPNYSRALMRVAAEYVEVTRVGDH 2100

QY 2101 YVTGMTTDNKKPCQVPAPEFFTEVDGVRLHRYAPACKPLLRDVTFOVGLNQYLVSQ L 2160

Db 2101 YVTGMTTDNKKPCQVPAPEFFTEVDGVRLHRYAPACKPLLRDVTFOVGLNQYLVSQ L 2160

QY 2161 PCEPEPDVTLTSMLTDPDSHITAE TAKRLARGSPPSLASSASQLSAPSLKATCTTHHD 2220

Db 2161 PCEPEPDVTLTSMLTDPDSHITAE TAKRLARGSPPSLASSASQLSAPSLKATCTTHHD 2220

QY 2221 SPDADLLEANLLWRQEMGNITRVESENKVILDSFEP LHAEGDERETSVA AEILRKSRK 2280

Db 2221 SPDADLLEANLLWRQEMGNITRVESENKVILDSFEP LHAEGDERETSVA AEILRKSRN 2280

QY 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESNV 2340

Db 2281 FPSALPIWARPDYNNPPLLESWKDPDYVPVHVHGCPLPPTKAPPIPPRRKRTVVLTESTV 2340

QY 2341 SSALAEALATKTFGSSGSSA VDSGTATALPD LASDDGKGSVDYESYSSNMPPLEGEGDPDL 2400

Db 2341 SSALAEALATKTFGSSGSSA VDSGTATAPPDQASDDGKESVDYESYSSNMPPLEGEGDPDL 2400

QY 2401 SDGSGSTVSEAS EADVCCSMYSYTWGALITPCA EESKLPINPLNSLLRHHNNVYATT 2460

Db 2401 SDGSGSTVSEATEDVCCSMYSYTWGALITPCA EESKLPINPLNSLLRHHNNVYATT 2460

QY 2461 SRSASLRQKVTDFRLQVLDHYRDVLKEMKAKASTVKAKLISIEBEACKLTPPHSAKSKF 2520

Db 2461 SRSASLRQKVTDFRLQVLDHYRDVLKEMKAKASTVKAKLISVBEACKLTPPHSAKSKF 2520

QY 2521 GYGAKDVRLSSRAVNHIRS VNEVDLLEDTETPIDTTIMAKSEVFCVQPEKGRKPARLIV 2580

Db 2521 GYGAKDVRLSSKAINHIRS VNKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIV 2580

Search completed: December 3, 2005, 08:27:30
Job time : 435 secs

QY 2581 FPDGLGVRVCEKMA LYDVVSTLPOAVMGSSVGYOYSPKORVEFLVNTWKSKCPMGFSYDT 2640

Db 2581 FPDGLGVRVCEKMA LYDVVSTLPOAVMGSSVGYOYSPKORVEFLVNTWKSKCPMGFSYDT 2640

QY 2641 RCFDSTVTWETSDIRVEESIYQCCCDLAP EAPQAIRSLTERLYIGPLTNSKGQNGCYRR CRA 2700

Db 2641 RCFDSTVTWETSDIRVEESIYQCCCDLAP EAPQAIRSLTERLYIGPLTNSKGQNGCYRR CRA 2700

QY 2701 SGVLTTS CGNTLT CYLKATAACRAAKLQDC TMLVNGDDLVI CESAGTQEDAAALRA FTE 2760

Db 2701 SGVLTTS CGNTLT CYLKATAACRAAKLQDC TMLVNGDDLVI CESAGTQEDAAALRA FTE 2760

QY 2761 AMTRY SAPPGDPPOPEYDLELITSCSSNVSAHDASGRVYVYLTROPTTTP LARA AAWETAR 2820

Db 2761 AMTRY SAPPGDPPOPEYDLELITSCSSNVSAHDASGRVYVYLTROPTTTP LARA AAWETAR 2820

QY 2821 HTPIN SWLGN I MYAP T L WARMILMTHFFS ILLAQEOL E KALDCQIYGACYSIEPLDLPQ 2880

Db 2821 HTPIN SWLGN I MYAP T L WARMILMTHFFS ILLAQEOL E KALDCQIYGACYSIEPLDLPQ 2880

QY 2881 IIERLHGLS AFTLHYSYSPGEINRVASCLRKLGVPPPLRTWHRARSVR AKLLSQGGRATC 2940

Db 2881 IIERLHGLS AFTLHYSYSPGEINRVASCLRKLGVPPPLRTWHRARSVR AKLLSQGGRATC 2940

QY 2941 GRYLFNNAVTRTKLKLTPIPAASQ L DLSGWFVAGYSGGDIYHSLSRARPRWFP LCLLLSV 3000

Db 2941 GRYLFNNAVTRTKLKLTPIPAASQ L DLSGWFVAGYSGGDIYHSLSRARPRWFP LCLLLSV 3000

QY 3001 GVGIYLLPNR 3010

Db 3001 GVGIYLLPNR 3010

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 04:27:12 ; Search time 30569 Seconds
(without alignments)
17842.027 Million cell updates/sec

Title: US-09-662-454-4
Perfect score: 9595
Sequence: 1 gccagccctgatgggggc.....ggcctctgcagatcatgt 9595

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 588141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_to.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	9595	100.0	9595	6 AR119832	AR119832 Sequence
2	9595	100.0	9595	6 CQ819761	CQ819761 Sequence
3	9595	100.0	9595	13 AF054247	AF054247 Hepatitis
4	9518.2	99.2	9595	13 AF054248	AF054248 Hepatitis
5	9473.6	98.7	9596	13 AF054249	AF054249 Hepatitis
6	9381.4	97.8	9580	13 AF054250	AF054250 Hepatitis
7	9327.4	97.2	9460	13 HPCJ491	D10750 Hepatitis C
8	9190.4	95.8	9448	13 HPCJ483	D13558 Hepatitis C
9	8780	91.5	8780	13 AF054255	AF054255 Hepatitis
10	8724	90.9	8780	13 AF054257	AF054257 Hepatitis
11	8703.2	90.7	8780	13 AF054252	AF054252 Hepatitis
12	8703.2	90.7	8780	13 AF054258	AF054258 Hepatitis
13	8684.8	90.5	8779	13 AF054251	AF054251 Hepatitis
14	8658.6	90.2	8781	13 AF054254	AF054254 Hepatitis
15	8624.6	89.9	9605	13 DQ071885	DQ071885 Hepatitis
16	8599.2	89.6	8780	13 AF054259	AF054259 Hepatitis
17	8592.8	89.6	8780	13 AF054253	AF054253 Hepatitis
18	8577	89.4	8781	13 AF054256	AF054256 Hepatitis

19	8502.6	88.6	9379	13 AF207758	AF207758 Hepatitis
20	8391.8	87.5	9435	13 AB049093	AB049093 Hepatitis
21	8387.2	87.4	9616	13 AB049088	AB049088 Hepatitis
22	8380.2	87.3	9587	13 AB080299	AB080299 Hepatitis
23	8357.4	87.1	9377	13 AF207756	AF207756 Hepatitis
24	8353.8	87.1	9547	13 AB049091	AB049091 Hepatitis
25	8348.2	87.0	9587	13 AF333324	AF333324 Hepatitis
26	8343.8	87.0	9573	13 AB049090	AB049090 Hepatitis
27	8342.8	86.9	9379	13 AF165059	AF165059 Hepatitis
28	8342.8	86.9	9379	13 AF165060	AF165060 Hepatitis
29	8339	86.9	9611	13 AY045702	AY045702 Hese sequen
30	8311.8	86.6	9413	6 E08399	E08399 Base sequen
31	8311.8	86.6	9413	6 E09631	E09631 HCV genome.
32	8311.8	86.6	9413	6 AR607373	AR607373 Sequence
33	8311.8	86.6	9413	13 HPCJCG	H98461 Smal-Rsr2 r
34	8302.2	86.5	9413	6 E08461	D90208 Hepatitis C
35	8299.2	86.5	9456	13 HPCRNA	D10934 Hepatitis C
36	8272.4	86.2	9379	13 AF165051	AF165051 Hepatitis
37	8271.6	86.2	9375	13 AF207754	AF207754 Hepatitis
38	8268.2	86.2	9379	13 AF207767	AF207767 Hepatitis
39	8262.4	86.1	9377	13 AF207763	AF207763 Hepatitis
40	8261	86.1	9374	13 AF207753	AF207753 Hepatitis
41	8250.2	86.0	9605	6 AX739971	AX739971 Sequence
42	8250.2	86.0	9605	13 HQJ238799	AJ238799 Hepatitis
43	8242.6	85.9	9379	13 AF165052	AF165052 Hepatitis
44	8242.6	85.9	9535	13 D85516	D85516 Hepatitis C
45	8230	85.8	9616	13 AF139594	AF139594 Hepatitis

ALIGNMENTS

RESULT 1
AR119832
LOCUS AR119832 9595 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 4 from patent US 6153421.
ACCESSION AR119832
VERSION AR119832.1 GI:14102531
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9595)
AUTHORS Yanagi,M., Bukh,J., Emerson,S.U. and Purcell,R.H.
TITLE Cloned genomes of infectious hepatitis C viruses and uses thereof
JOURNAL Patent: US 6153421-A 4 28-NOV-2000;
FEATURES
source 1..9595
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 9595; DB 6; Length 9595;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATACTG 60
Db 1 GCCAGCCCCCTGATGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATACTG 60
Qy 61 TCTTACGCGAGAAAGCCTTAGCCATGGCGTTAGTATGATGTCGTGAGCTCCAGGAC 120
Db 61 TCTTACGCGAGAAAGCCTTAGCCATGGCGTTAGTATGATGTCGTGAGCTCCAGGAC 120
Qy 121 CCCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180
Db 121 CCCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180
Qy 181 GACGACCGGGTCTTTCTTTGGATCAACCCGCTCAATGCTCGAGATTGGCGTGCCTCC 240
Db 181 GACGACCGGGTCTTTCTTTGGATCAACCCGCTCAATGCTCGAGATTGGCGTGCCTCC 240
Qy 241 GCGAGACTGTAGCCGAGTAGTGTGGGTCCGAAAGCCCTTGTGGTACTGCCTGATAGG 300

[illegible]

QY	2461	CAGCGTTTGTCTCTTTGCAATCAAAATGGAGTACATCCCTGTGCTTTTCCCTTCTCCTGG	2520
DB	2461	CAGCGTTTGTCTCTTTGCAATCAAAATGGAGTACATCCCTGTGCTTTTCCCTTCTCCTGG	2520
QY	2521	CAGACGGCGGTGCTGCTGCTGCTGATGATGCTGCTGATAGCCAGGCTAGGCGG	2580
DB	2521	CAGACGGCGGTGCTGCTGCTGCTGATGATGCTGCTGATAGCCAGGCTAGGCGG	2580
QY	2581	CCTTAGAGAACTTGGTGTCTCAATGCGCGCTCCGTGGCGGAGCGCATGTATTCTCT	2640
DB	2581	CCTTAGAGAACTTGGTGTCTCAATGCGCGCTCCGTGGCGGAGCGCATGTATTCTCT	2640
QY	2641	CCTTTCTTGTGTTCTTCTGCGCGCTGTGATTAAGGCGAGGCTGCTCTTGGGCGG	2700
DB	2641	CCTTTCTTGTGTTCTTCTGCGCGCTGTGATTAAGGCGAGGCTGCTCTTGGGCGG	2700
QY	2701	CGTATGCTTTTATGGGTATGGCGCTGCTCTGCTCTCTACTGGCGTTACCAACGAG	2760
DB	2701	CGTATGCTTTTATGGGTATGGCGCTGCTCTGCTCTCTACTGGCGTTACCAACGAG	2760
QY	2761	CTTACGCTTGGACCGGAGATGGCTGCATCGTCGGGGGTGCGGTTCTTGTAGGCTGG	2820
DB	2761	CTTACGCTTGGACCGGAGATGGCTGCATCGTCGGGGGTGCGGTTCTTGTAGGCTGG	2820
QY	2821	TATTCTTGACCTTGTCAACCATCTACAAAGTGTTCCTACTAGGCTCATATGTGTTAC	2880
DB	2821	TATTCTTGACCTTGTCAACCATCTACAAAGTGTTCCTACTAGGCTCATATGTGTTAC	2880
QY	2881	AATACTTTATCAGAGCGGCGGCACATGCAAGTGTGGTCCCGCCCTCAAGTTC	2940
DB	2881	AATACTTTATCAGAGCGGCGGCACATGCAAGTGTGGTCCCGCCCTCAAGTTC	2940
QY	2941	GGGAGGCGCGATGCCATCATCTCTCCTCAGCTGTGCGGTTCTACAGAGTTAATTTTG	3000
DB	2941	GGGAGGCGCGATGCCATCATCTCTCCTCAGCTGTGCGGTTCTACAGAGTTAATTTTG	3000
QY	3001	ACATCAACAAACTCTGCTCGCCATCTCGGCCCGCTCATGGTGTCTCCAGGCTGGCATAA	3060
DB	3001	ACATCAACAAACTCTGCTCGCCATCTCGGCCCGCTCATGGTGTCTCCAGGCTGGCATAA	3060
QY	3061	CGAGAGTCCGCTCTCGTGGCGCTCAAGGCTCATCTGTCATGATGATGTAGTCGAA	3120
DB	3061	CGAGAGTCCGCTCTCGTGGCGCTCAAGGCTCATCTGTCATGATGATGTAGTCGAA	3120
QY	3121	AAGTCGCGGGGTCAATTATGTCCAAATGTCTTCATGAAGCTGGGCGCTGACAGGTA	3180
DB	3121	AAGTCGCGGGGTCAATTATGTCCAAATGTCTTCATGAAGCTGGGCGCTGACAGGTA	3180
QY	3181	CGTACGTTTATTAACCATCTTACCCACTGCGGGACTGGGCCACGCGGGCTACGAGACC	3240
DB	3181	CGTACGTTTATTAACCATCTTACCCACTGCGGGACTGGGCCACGCGGGCTACGAGACC	3240
QY	3241	TTGGGTGGGGTATAGACCGCTGCTCTTCTCCGCTATGGAGCAAGGCTCATCCTGGG	3300
DB	3241	TTGGGTGGGGTATAGACCGCTGCTCTTCTCCGCTATGGAGCAAGGCTCATCCTGGG	3300
QY	3301	GAGCAGACACGCTGCTGGGACATCATCTTTGGGTCTACCCGCTCCGCCGGAAGGG	3360
DB	3301	GAGCAGACACGCTGCTGGGACATCATCTTTGGGTCTACCCGCTCCGCCGGAAGGG	3360
QY	3361	GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTTGGCGACTCCTTG	3420
DB	3361	GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTTGGCGACTCCTTG	3420
QY	3421	CGCCCATCAAGGCTTCTCCAAACGCGGGGCTACTTGGTTGATCATCATAGACC	3480
DB	3421	CGCCCATCAAGGCTTCTCCAAACGCGGGGCTACTTGGTTGATCATCATAGACC	3480
QY	3481	TCACAGCGGGGACAAAGACGAGTTCGAAAGGGAGGTTCAAGTGTTCCTACCCGACAC	3540
DB	3481	TCACAGCGGGGACAAAGACGAGTTCGAAAGGGAGGTTCAAGTGTTCCTACCCGACAC	3540

QY	3541	AATCTTCTCGGACCTGCATCAAGCGGCTGTGTGGACTGTCTACCATGGCGCTGGCT	3600
DB	3541	AATCTTCTCGGACCTGCATCAAGCGGCTGTGTGGACTGTCTACCATGGCGCTGGCT	3600
QY	3601	CGAAGACCTAGCCGCTCCAAAAGGTCCAAATCACCCAAATGTACACCAATGTAGACCTGG	3660
DB	3601	CGAAGACCTAGCCGCTCCAAAAGGTCCAAATCACCCAAATGTACACCAATGTAGACCTGG	3660
QY	3661	ACTCTGCTGGCTGAGGCGCCCCCGGGCGGCTCCATGACACCATGACGCTGTGGCA	3720
DB	3661	ACTCTGCTGGCTGAGGCGCCCCCGGGCGGCTCCATGACACCATGACGCTGTGGCA	3720
QY	3721	GCTCGGACCTTTACTTTGGTCAAGACATGCTGATGCTCATTCGCGTGGCGCGGAGGCG	3780
DB	3721	GCTCGGACCTTTACTTTGGTCAAGACATGCTGATGCTCATTCGCGTGGCGCGGAGGCG	3780
QY	3781	ACAGCAGGGAAGTCTACTCTCCCCAGGCCCGTCTCTACTGAAAGGCTCTCTCGGGTG	3840
DB	3781	ACAGCAGGGAAGTCTACTCTCCCCAGGCCCGTCTCTACTGAAAGGCTCTCTCGGGTG	3840
QY	3841	GTCCATTTGCTTTCGCGGACGCTGTGGGCGCTCTTCCGGGCTGCTGTGTGACACC	3900
DB	3841	GTCCATTTGCTTTCGCGGACGCTGTGGGCGCTCTTCCGGGCTGCTGTGTGACACC	3900
QY	3901	GGGGGTGCGAAGGCGGTGGACTTCATACCCGTTGAGTCTATGGAACCTACCATCGGT	3960
DB	3901	GGGGGTGCGAAGGCGGTGGACTTCATACCCGTTGAGTCTATGGAACCTACCATCGGT	3960
QY	3961	CTCCGCTCTTCAAGCAAACTCAACCCCCCGGCTGTACCGCAGACATTCCAAAGTGGCAC	4020
DB	3961	CTCCGCTCTTCAAGCAAACTCAACCCCCCGGCTGTACCGCAGACATTCCAAAGTGGCAC	4020
QY	4021	ATCTGACGCTCTACTGCGAGCGGCAAGACACAAAGTCCGGGCTGCTGTGTGACGCC	4080
DB	4021	ATCTGACGCTCTACTGCGAGCGGCAAGACACAAAGTCCGGGCTGCTGTGTGACGCC	4080
QY	4081	AAGGTACAGGTGCTGCTGAAACCGCTGCTGCGCCACTTAGGGTTTGGGCGCT	4140
DB	4081	AAGGTACAGGTGCTGCTGAAACCGCTGCTGCGCCACTTAGGGTTTGGGCGCT	4140
QY	4141	ATATGTCCAAGGCACACGCTATCGACCTTAAATCAGAACTGGGGTAAGGACCATTAACA	4200
DB	4141	ATATGTCCAAGGCACACGCTATCGACCTTAAATCAGAACTGGGGTAAGGACCATTAACA	4200
QY	4201	CGGCGGCTCCATTACGTACTCAACCTATGGCAAGTTCCTTGGCGACGCTGGCTGTCTG	4260
DB	4201	CGGCGGCTCCATTACGTACTCAACCTATGGCAAGTTCCTTGGCGACGCTGGCTGTCTG	4260
QY	4261	GGGGCGCTATGACATCATTAATATGTGATGAGTGCCACTCAACTGCTACCATCTCT	4320
DB	4261	GGGGCGCTATGACATCATTAATATGTGATGAGTGCCACTCAACTGCTACCATCTCT	4320
QY	4321	TGGGCATCGGCACAGTCTCGGACCAAGCGAGACGCTGGAGCGGCTCGTCTGTCTG	4380
DB	4321	TGGGCATCGGCACAGTCTCGGACCAAGCGAGACGCTGGAGCGGCTCGTCTGTCTG	4380
QY	4381	CCACCGCTACACCTCCGGGATCGGTTACCGTGGCCACACCCCAATATCGAGAAATAGGCC	4440
DB	4381	CCACCGCTACACCTCCGGGATCGGTTACCGTGGCCACACCCCAATATCGAGAAATAGGCC	4440
QY	4441	TGTCCAAATGAGAGATCCCTTCTATGGCAAGGCATCCCATTTGAGGCCATCAAGG	4500
DB	4441	TGTCCAAATGAGAGATCCCTTCTATGGCAAGGCATCCCATTTGAGGCCATCAAGG	4500
QY	4501	GGGGGAGGCTCTCATTTTCTGCCATTCGAAGAAATGTGACGAGCTCGCCGCAAGC	4560
DB	4501	GGGGGAGGCTCTCATTTTCTGCCATTCGAAGAAATGTGACGAGCTCGCCGCAAGC	4560
QY	4561	TGACAGGCTCGGACTGAAAGGCTGTAGCATATTTACCGGGGCTTGTGTGTCTGTATAC	4620
DB	4561	TGACAGGCTCGGACTGAAAGGCTGTAGCATATTTACCGGGGCTTGTGTGTCTGTATAC	4620
QY	4621	CGCTATCGGAGACGCTCGTTGTCTGTGGCAACAGACGCTCTTAATGACGGGTTTACCGGCG	4680

Db 4621 |||||CGCCTATCGGAGACGTGTGTGCTGGCAACAGACGCTCTAATGACGGGTTTCACCGCGC 4680
Qy 4681 ATTTTGGACTCAGTGATCGACTGCAATACATGTGTCCACCAGACAGTGCACATTCAGCTTGG 4740
Db 4681 ATTTTGGACTCAGTGATCGACTGCAATACATGTGTCCACCAGACAGTGCACATTCAGCTTGG 4740
Qy 4741 ATCCCACTTACACATTTGAGACGACGACGCGTCCCAAGACGCGGTGTGCGCTCGCAAC 4800
Db 4741 ATCCCACTTACACATTTGAGACGACGACGCGTCCCAAGACGCGGTGTGCGCTCGCAAC 4800
Qy 4801 GCGAGGTAGAACTGCGAGGGTAGGATGGGCATCTACAGTTTGTGTACTCCAGAGAAC 4860
Db 4801 GCGAGGTAGAACTGCGAGGGTAGGATGGGCATCTACAGTTTGTGTACTCCAGAGAAC 4860
Qy 4861 GCGCCTCGGGCATGTTCGATTTCTCGGTCTGTGTGAGTGTATGACGCGGGCTGTGCTT 4920
Db 4861 GCGCCTCGGGCATGTTCGATTTCTCGGTCTGTGTGAGTGTATGACGCGGGCTGTGCTT 4920
Qy 4921 GGTATGAGCTCACGCCGCTGAGACTCGTGTAGGTTCGGGCTTACCTAAATACACCAG 4980
Db 4921 GGTATGAGCTCACGCCGCTGAGACTCGTGTAGGTTCGGGCTTACCTAAATACACCAG 4980
Qy 4981 GGTGCCCTGTGCGCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTCAAGSCCTCACCC 5040
Db 4981 GGTGCCCTGTGCGCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTCAAGSCCTCACCC 5040
Qy 5041 ACATAGATGCCACATTCCTGTCTCCAGACTAAACAGGCAAGGAGACAACTTCTTACCTGG 5100
Db 5041 ACATAGATGCCACATTCCTGTCTCCAGACTAAACAGGCAAGGAGACAACTTCTTACCTGG 5100
Qy 5101 TGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCATCTGTGGGACCAAA 5160
Db 5101 TGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCATCTGTGGGACCAAA 5160
Qy 5161 TGTGGAAGTGTCTCATACGGCTGAAACCTTACACTGACGCGGCCAACACCCCTGCTGTATA 5220
Db 5161 TGTGGAAGTGTCTCATACGGCTGAAACCTTACACTGACGCGGCCAACACCCCTGCTGTATA 5220
Qy 5221 GCGTAGGAGCCGTCCAAAATGAGGTCTCCTCACACACCCCAATAACATCATG 5280
Db 5221 GCGTAGGAGCCGTCCAAAATGAGGTCTCCTCACACACCCCAATAACATCATG 5280
Qy 5281 CATGATGTGCGCTGACCTGGAGGTGCTCACTAGACACCTGGTGTGTAGGCGAGTCC 5340
Db 5281 CATGATGTGCGCTGACCTGGAGGTGCTCACTAGACACCTGGTGTGTAGGCGAGTCC 5340
Qy 5341 TTGCACTTTGGCCGCGATCTGCTGACGACAGGAGTGTGTCTATTGTGGSCAGGATCA 5400
Db 5341 TTGCACTTTGGCCGCGATCTGCTGACGACAGGAGTGTGTCTATTGTGGSCAGGATCA 5400
Qy 5401 TCTTGTCCGGGAAGCCAGCTGTCTGTTCCCGACAGGGAAGTCTCTACCAAGGAGTTCGATG 5460
Db 5401 TCTTGTCCGGGAAGCCAGCTGTCTGTTCCCGACAGGGAAGTCTCTACCAAGGAGTTCGATG 5460
Qy 5461 AGATGAAGAGTGTGCTTCACTTCACTGAGCAGGGAATGACAGTTCGCGGAGC 5520
Db 5461 AGATGAAGAGTGTGCTTCACTTCACTGAGCAGGGAATGACAGTTCGCGGAGC 5520
Qy 5521 AATTCAAGCAAAAGGCGCTCGGTTGTGCAAAACGGCCCAAGCAAGCGGAGGCTGCTG 5580
Db 5521 AATTCAAGCAAAAGGCGCTCGGTTGTGCAAAACGGCCCAAGCAAGCGGAGGCTGCTG 5580
Qy 5581 CTCCTGTGTGAGTCCAAAGTGCAGCCCTTGAGACCTTCTGGGGAAGCAATGTGGA 5640
Db 5581 CTCCTGTGTGAGTCCAAAGTGCAGCCCTTGAGACCTTCTGGGGAAGCAATGTGGA 5640
Qy 5641 ATTTCAATACGCGAATACAGTACCTAGCAGGCTTATCCACTGTGCTGGAACCCCGGA 5700
Db 5641 ATTTCAATACGCGAATACAGTACCTAGCAGGCTTATCCACTGTGCTGGAACCCCGGA 5700
Qy 5701 TAGCATCATTTGATGGCATTTTACAGCTTCTATCACTAGCCCGCTTCACCAACCCCAACCC 5760
Db 5701 TAGCATCATTTGATGGCATTTTACAGCTTCTATCACTAGCCCGCTTCACCAACCCCAACCC 5760

Db 5701 TAGCATCATTTGATGGCATTTTACAGCTTCTATCACTAGCCCGCTTCACCAACCCCAACCC 5760
Qy 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGCTGCCAACTCGCTCCCTCCAGCGCTCGT 5820
Db 5761 TCCTGTTTAAACATCTTGGGGGATGGGTGCTGCCAACTCGCTCCCTCCAGCGCTCGT 5820
Qy 5821 CAGCTTTCTGTTGGCGCGCATCGCGGAGCGGTGTGTGGCAGCATAGGCTTGGGAAG 5880
Db 5821 CAGCTTTCTGTTGGCGCGCATCGCGGAGCGGTGTGTGGCAGCATAGGCTTGGGAAG 5880
Qy 5881 TGCTGTGGACATCTTGGCGGGCTATGGGCGAGGGTAGCGGGCGCACTCGTGGCCCTTTA 5940
Db 5881 TGCTGTGGACATCTTGGCGGGCTATGGGCGAGGGTAGCGGGCGCACTCGTGGCCCTTTA 5940
Qy 5941 AGGTCATAGCGGCGAGGTGCCCTCCACGAGGACCTGGTCAACTTACTTCCCTCCCATCC 6000
Db 5941 AGGTCATAGCGGCGAGGTGCCCTCCACGAGGACCTGGTCAACTTACTTCCCTCCCATCC 6000
Qy 6001 TCTCTCTGTGTCCTCGTCTCGGGGTCTGTGCGAGCAATACTCGTCCGACAGTGG 6060
Db 6001 TCTCTCTGTGTCCTCGTCTCGGGGTCTGTGCGAGCAATACTCGTCCGACAGTGG 6060
Qy 6061 GCCCGGAGAGGGGCTGTGAGTGAACCGGCTGATAGGCTTCGCTTCGCGGGGTA 6120
Db 6061 GCCCGGAGAGGGGCTGTGAGTGAACCGGCTGATAGGCTTCGCTTCGCGGGGTA 6120
Qy 6121 ACCAGCTCTCCCTTACGACATCTGTGCTGAGAGCAGCTGACAGCAGTGTCTCAGA 6180
Db 6121 ACCAGCTCTCCCTTACGACATCTGTGCTGAGAGCAGCTGACAGCAGTGTCTCAGA 6180
Qy 6181 TCTCTCTAGCTTTACCATCACTCACTGCTGAAGCGGCTCCACCACTGAGTAAATGAGG 6240
Db 6181 TCTCTCTAGCTTTACCATCACTCACTGCTGAAGCGGCTCCACCACTGAGTAAATGAGG 6240
Qy 6241 ACTGCTCTACGCAATGCTCGGCTCGTGAAGGATGTTTGGATTTGATGACAGG 6300
Db 6241 ACTGCTCTACGCAATGCTCGGCTCGTGAAGGATGTTTGGATTTGATGACAGG 6300
Qy 6301 TGTGACTGACTTCAAGACCTTGGCTCGAGTCCAACTCTCGCGCGGTTTACCGGAGTCC 6360
Db 6301 TGTGACTGACTTCAAGACCTTGGCTCGAGTCCAACTCTCGCGCGGTTTACCGGAGTCC 6360
Qy 6361 CTTTCTCTGTCATGCAACCGGGTACAAGGAGTCTGCGCGGGGAGCGCATCATGCAAA 6420
Db 6361 CTTTCTCTGTCATGCAACCGGGTACAAGGAGTCTGCGGGGAGCGCATCATGCAAA 6420
Qy 6421 CCACTGCGCCATGCGAGACAGATCGCGGACATGTCAAAAACGTTTCCATGAGGATCG 6480
Db 6421 CCACTGCGCCATGCGAGACAGATCGCGGACATGTCAAAAACGTTTCCATGAGGATCG 6480
Qy 6481 TAGGCGCTTAGAACCTGACGACACACGTCGCAACGTTTCCCATCAACGATACACCA 6540
Db 6481 TAGGCGCTTAGAACCTGACGACACACGTCGCAACGTTTCCCATCAACGATACACCA 6540
Qy 6541 CCGGACCTTGCACACCTTCCCGCGGCCCAACTATTCCAGGGCGCTATGGCGGTGGCTG 6600
Db 6541 CCGGACCTTGCACACCTTCCCGCGGCCCAACTATTCCAGGGCGCTATGGCGGTGGCTG 6600
Qy 6601 CTGAGGAGTACGTTGAGGTTACGCGTGTGGGGATTTCCACTAGCTGACGGGCGATGACCA 6660
Db 6601 CTGAGGAGTACGTTGAGGTTACGCGTGTGGGGATTTCCACTAGCTGACGGGCGATGACCA 6660
Qy 6661 CTGACAAAGTAAAGTCCCATGCGGTTCCGGCCCCCGAATTTCTTACGGAGGTGGATG 6720
Db 6661 CTGACAAAGTAAAGTCCCATGCGGTTCCGGCCCCCGAATTTCTTACGGAGGTGGATG 6720
Qy 6721 GAGTGCCTTGCACAGGTACGCTCCGGCGTCAAAACCTTCTTACGGGAGGACGTCAGT 6780
Db 6721 GAGTGCCTTGCACAGGTACGCTCCGGCGTCAAAACCTTCTTACGGGAGGACGTCAGT 6780
Qy 6781 TCCAGGTGCGGCTCAACCAATACCTTGTGTCGGGTGCGAGCTCCCATGCGAGCCGACCCG 6840
Db 6781 TCCAGGTGCGGCTCAACCAATACCTTGTGTCGGGTGCGAGCTCCCATGCGAGCCGACCCG 6840

QY	6841	ACGTAAACAGTGCCTTATTCATAGCTCAACCGATCCCTCCCAATTAAGAGAGACGGCTA	6900	QY	7921	TCCGGAACCTATCCAGCAGGCGGTTAAACCATCCGCTCCGTGTGGGAGGACTTCTCTGG	7980
Db	6841	ACGTAAACAGTGCCTTATTCATAGCTCAACCGATCCCTCCCAATTAAGAGAGACGGCTA	6900	Db	7921	TCCGGAACCTATCCAGCAGGCGGTTAAACCATCCGCTCCGTGTGGGAGGACTTCTCTGG	7980
QY	6901	ACGCTAGGCTGGCTAGAGGGTCTCCCCCTCTTTTAGCCAGTCTCATCAGCTTAGCCAGTTGT	6960	QY	7981	AAGACACTGAAACACCAATTGACACCAATCATGCGCAAAAAGTGAAGGTTTCTGCGTCC	8040
Db	6901	ACGCTAGGCTGGCTAGAGGGTCTCCCCCTCTTTTAGCCAGTCTCATCAGCTTAGCCAGTTGT	6960	Db	7981	AAGACACTGAAACACCAATTGACACCAATCATGCGCAAAAAGTGAAGGTTTCTGCGTCC	8040
QY	6961	CTGCGCCTCTTTTGAAGGCGACATGCACTACCAACCATGACTCCCGGACGCTGACTCA	7020	QY	8041	AACCAGAGAAGGAGCGCGCAAGCCAGCTCGCCTTATCCCTATTCAGACCTCGGAGTTTC	8100
Db	6961	CTGCGCCTCTTTTGAAGGCGACATGCACTACCAACCATGACTCCCGGACGCTGACTCA	7020	Db	8041	AACCAGAGAAGGAGCGCGCAAGCCAGCTCGCCTTATCCCTATTCAGACCTCGGAGTTTC	8100
QY	7021	TCGAGGCAACCTCTTTGTGGCGGAGAGATGGCGGGAACATCACTCGCGTGGAGTCAG	7080	QY	8101	GTGTATGCGAGAGAATGCGCCCTTTACGAGTGTCTCCACCTTCTCAGGCCGTGATGG	8160
Db	7021	TCGAGGCAACCTCTTTGTGGCGGAGAGATGGCGGGAACATCACTCGCGTGGAGTCAG	7080	Db	8101	GTGTATGCGAGAGAATGCGCCCTTTACGAGTGTCTCCACCTTCTCAGGCCGTGATGG	8160
QY	7081	AGAAATAGGTAGTAATTTCTCGACTCTTTTCGAAACCGCTTTCACGCGAGGGGATGAGAGG	7140	QY	8161	GCTCTCTCATACGGATTTCAATCTCCCAAGCAGCGGGTCGAGTTCTCTGGTGAATACCT	8220
Db	7081	AGAAATAGGTAGTAATTTCTCGACTCTTTTCGAAACCGCTTTCACGCGAGGGGATGAGAGG	7140	Db	8161	GCTCTCTCATACGGATTTCAATCTCCCAAGCAGCGGGTCGAGTTCTCTGGTGAATACCT	8220
QY	7141	AGATATCCGTGCGGCGGAGATCTCGGAAAATCCAGGAAATTCCTCAGCGTTGCCCCA	7200	QY	8221	GGAAATCAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGTGACTCAACGG	8280
Db	7141	AGATATCCGTGCGGCGGAGATCTCGGAAAATCCAGGAAATTCCTCAGCGTTGCCCCA	7200	Db	8221	GGAAATCAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGTGACTCAACGG	8280
QY	7201	TATGGGCAACCCCGGACTACAATCTCTCCACTGTGTAGAGTCTCTGGAAGGACCCGACTACG	7260	QY	8281	TCACTGAGAGTGCATTTCTGTGTTGAGAGTCAATTTACCAATGTTGTGACTTGGCCCCCG	8340
Db	7201	TATGGGCAACCCCGGACTACAATCTCTCCACTGTGTAGAGTCTCTGGAAGGACCCGACTACG	7260	Db	8281	TCACTGAGAGTGCATTTCTGTGTTGAGAGTCAATTTACCAATGTTGTGACTTGGCCCCCG	8340
QY	7261	TCCCTCCGGTGTACAGGATGCCATTTGCCACTACCAAGGCTCTCCCAATACCACTC	7320	QY	8341	AGCCACAGACAGGCGCATTAAGTCTCACAGAGCGGCTTTACATCGGGGTCCTCTGACTA	8400
Db	7261	TCCCTCCGGTGTACAGGATGCCATTTGCCACTACCAAGGCTCTCCCAATACCACTC	7320	Db	8341	AGCCACAGACAGGCGCATTAAGTCTCACAGAGCGGCTTTACATCGGGGTCCTCTGACTA	8400
QY	7321	CACGAGAAAGAGACGGTTGTCTCTGACAGAAATCCAAATGTTCTTCTGCTTGGCGGAGC	7380	QY	8401	ACTCAAAAAGGCGAGAACTCGCGGTTATCGCGGTCGCGCGCAAGTGGGCTCTGACGACTA	8460
Db	7321	CACGAGAAAGAGACGGTTGTCTCTGACAGAAATCCAAATGTTCTTCTGCTTGGCGGAGC	7380	Db	8401	ACTCAAAAAGGCGAGAACTCGCGGTTATCGCGGTCGCGCGCAAGTGGGCTCTGACGACTA	8460
QY	7381	TCGSCACTAAGACCTTCGGTAGCTCCGGATCGTCCGCCGTTGATAGCGGACCGCGACCG	7440	QY	8461	GCTCGCGTAAATACCTTCACATGTTACTTTGAAGCCACTGTCAGAGCTGTCGAGCTGCAAGC	8520
Db	7381	TCGSCACTAAGACCTTCGGTAGCTCCGGATCGTCCGCCGTTGATAGCGGACCGCGACCG	7440	Db	8461	GCTCGCGTAAATACCTTCACATGTTACTTTGAAGCCACTGTCAGAGCTGTCGAGCTGCAAGC	8520
QY	7441	CCCTTCTGTAGCTGGCTCCGACGACGGTGAACAAAGATCCGAGCTTGAGTCTGACTCCT	7500	QY	8521	TCAGAGACTGCACGATGCTCGTGAAACCGGACGACCTTGTGCTTATCTGTGAAAGCGCG	8580
Db	7441	CCCTTCTGTAGCTGGCTCCGACGACGGTGAACAAAGATCCGAGCTTGAGTCTGACTCCT	7500	Db	8521	TCAGAGACTGCACGATGCTCGTGAAACCGGACGACCTTGTGCTTATCTGTGAAAGCGCG	8580
QY	7501	CCATGCCCCCTTTGAAGGGAGCCGGGACCCCGATCTCAGCGACGGGCTTTGGTCTA	7560	QY	8581	GAACCCAGGAGGATGCGCGCGCTTACGAGCCTTTCACGAGGCTATGACTAGTATTCGG	8640
Db	7501	CCATGCCCCCTTTGAAGGGAGCCGGGAGCCCGATCTCAGCGACGGGCTTTGGTCTA	7560	Db	8581	GAACCCAGGAGGATGCGCGCGCTTACGAGCCTTTCACGAGGCTATGACTAGTATTCGG	8640
QY	7561	CCGTGAGTGAAGGAGCTAGTGAGGATGCTGCTGCTCAATGTCCTTATACGTGGACAG	7620	QY	8641	CCCCCCCCGGGATCCGCCCAACCGAGATACGACCTGGAGCTGATAACATCATGTTCTCT	8700
Db	7561	CCGTGAGTGAAGGAGCTAGTGAGGATGCTGCTGCTCAATGTCCTTATACGTGGACAG	7620	Db	8641	CCCCCCCCGGGATCCGCCCAACCGAGATACGACCTGGAGCTGATAACATCATGTTCTCT	8700
QY	7621	CGCCCTGTATCAGCATGCGCTCGGAGGAAAGTAAAGTCCCATCAACCCGTTGAGCA	7680	QY	8701	CCAAATGTGTCAAGTGCAGCATGCTGCGCAAAAGGGTATACCTCACCCCGTGACC	8760
Db	7621	CGCCCTGTATCAGCATGCGCTCGGAGGAAAGTAAAGTCCCATCAACCCGTTGAGCA	7680	Db	8701	CCAAATGTGTCAAGTGCAGCATGCTGCGCAAAAGGGTATACCTCACCCCGTGACC	8760
QY	7681	ACTCTTTCTCGGTACCAACATGTTCTACGCCACAACATCCCGACGCAAGCCCTCC	7740	QY	8761	CCACCACCCCTTTGACGGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAACTCTT	8820
Db	7681	ACTCTTTCTCGGTACCAACATGTTCTACGCCACAACATCCCGACGCAAGCCCTCC	7740	Db	8761	CCACCACCCCTTTGACGGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAACTCTT	8820
QY	7741	GGCAGAAAGGTACCTTTGACAGATGCAAGTCTCGATGATCATTTACCGGACGCTAC	7800	QY	8821	GGCTAGGCAATATCATGTATGCGGCCCACTTATGGGCAAGGATGATTTCTGATGACTC	8880
Db	7741	GGCAGAAAGGTACCTTTGACAGATGCAAGTCTCGATGATCATTTACCGGACGCTAC	7800	Db	8821	GGCTAGGCAATATCATGTATGCGGCCCACTTATGGGCAAGGATGATTTCTGATGACTC	8880
QY	7801	TCAAGGAGATGAAGCGAAGCGTCCAAGTTAAGGCTTAAGCTTCTATAGAGGAGG	7860	QY	8881	ACTTTTCTCCATCTTCTAGCTCAAGAGCACTTTGAAAAGCCCTGGATGTGTCAGATCT	8940
Db	7801	TCAAGGAGATGAAGCGAAGCGTCCAAGTTAAGGCTTAAGCTTCTATAGAGGAGG	7860	Db	8881	ACTTTTCTCCATCTTCTAGCTCAAGAGCACTTTGAAAAGCCCTGGATGTGTCAGATCT	8940
QY	7861	CCTGCAAGCTGACGCCCCCACTTCGGGCAATCCAAATTTGGCTATCGGCGCAAGGAGC	7920	QY	8941	ACGGGGCTTGCTACTCCTCAATTGAGCCACTTGACCTCTCAGATCATTTGAACGACTCCATG	9000
Db	7861	CCTGCAAGCTGACGCCCCCACTTCGGGCAATCCAAATTTGGCTATCGGCGCAAGGAGC	7920	Db	8941	ACGGGGCTTGCTACTCCTCAATTGAGCCACTTGACCTCTCAGATCATTTGAACGACTCCATG	9000
				QY	9001	GTCTTAGGCAATTTAGACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060

[illegible]

Qy	1201	TCTCCAGCTGTTTCACTTCTCGCCTCGCGGCATGAGACAGTGCAGGACTGCAACTGCT	1260
Db	1201	TCTCCAGCTGTTTCACTTCTCGCCTCGCGGCATGAGACAGTGCAGGACTGCAACTGCT	1260
Qy	1261	CAATCTATCCCGGCATGATATCAGGTCAACCGATGGCTTGGGATATGATGAACTGGT	1320
Db	1261	CAATCTATCCCGGCATGATATCAGGTCAACCGATGGCTTGGGATATGATGAACTGGT	1320
Qy	1321	CACCTAACAGACCCCTAGTGGTGTCCAGTTGCTCCGGATCCCACAAAGCTGTCTGGACAC	1380
Db	1321	CACCTAACAGACCCCTAGTGGTGTCCAGTTGCTCCGGATCCCACAAAGCTGTCTGGACAC	1380
Qy	1381	TGTTGGCGGGGCCCACTGGCGAGTCTGTGGCGGCCCTTGCTACTATTCCATGGTAGGA	1440
Db	1381	TGTTGGCGGGGCCCACTGGCGAGTCTGTGGCGGCCCTTGCTACTATTCCATGGTAGGA	1440
Qy	1441	ACTGGGCTAAGTTTCTGAATTTGGCGCTACTCTTTTGC CGCGCTTTGACGGGAGACCCACA	1500
Db	1441	ACTGGGCTAAGTTTCTGAATTTGGCGCTACTCTTTTGC CGCGCTTTGACGGGAGACCCACA	1500
Qy	1501	CGACGGGAGGTTGGCGGCCACACACCTTC CGGTTACGTTCCCTTTTCTCATCTGGGG	1560
Db	1501	CGACGGGAGGTTGGCGGCCACACACCTTC CGGTTACGTTCCCTTTTCTCATCTGGGG	1560
Qy	1561	CGTCTCAGAAAAATCCAGCTTGTGAATACCAACCGCAGCTGGCACATCAACAGGACTGCC	1620
Db	1561	CGTCTCAGAAAAATCCAGCTTGTGAATACCAACCGCAGCTGGCACATCAACAGGACTGCC	1620
Qy	1621	TAAATTCGAATGACTCCCTCCAACTGGGTTCTTTTGC CGCGCTTTTACGCACACAAAGT	1680
Db	1621	TAAATTCGAATGACTCCCTCCAACTGGGTTCTTTTGC CGCGCTTTTACGCACACAAAGT	1680
Qy	1681	TCAACTCGTCCGGTGC CGGAGCGCATGGCAGCTGCCGCCCATTTAGCTGGTTCGCC	1740
Db	1681	TCAACTCGTCCGGTGC CGGAGCGCATGGCAGCTGCCGCCCATTTAGCTGGTTCGCC	1740
Qy	1741	AGGGTGGGGCCCATCACTATTAAGCTTAACAGCTCGGATCAGAGGCTTATTGCT	1800
Db	1741	AGGGTGGGGCCCATCACTATTAAGCTTAACAGCTCGGATCAGAGGCTTATTGCT	1800
Qy	1801	GGCATTACGGCCTCGACCGTGTGTTGTTGGGGAACACCGATCGTTCCGTTCCCTACGT	1860
Db	1801	GGCATTACGGCCTCGACCGTGTGTTGTTGGGGAACACCGATCGTTCCGTTCCCTACGT	1860
Qy	1861	ATTGTTTACCCCAAGCCCTTGTGTTGGGGAACACCGATCGTTCCGTTCCCTACGT	1920
Db	1861	ATTGTTTACCCCAAGCCCTTGTGTTGGGGAACACCGATCGTTCCGTTCCCTACGT	1920
Qy	1921	ATAGCTGGGGAGAAATGAGACAGCTGATGCTCTCAACACACCGCTCCGCCACAAG	1980
Db	1921	ATAGCTGGGGAGAAATGAGACAGCTGATGCTCTCAACACACCGCTCCGCCACAAG	1980
Qy	1981	GCAACTGTTCCGCTGTACATGGATGAATAGTACTTGGGTTTCACTAAGACGTGGAGGTC	2040
Db	1981	GCAACTGTTCCGCTGTACATGGATGAATAGTACTTGGGTTTCACTAAGACGTGGAGGTC	2040
Qy	2041	CCCGTGTAAACATCGGGGGGTTCGGTAAACCGACCTTGAATCTGCCCAACCGACTGCTTCC	2100
Db	2041	CCCGTGTAAACATCGGGGGGTTCGGTAAACCGACCTTGAATCTGCCCAACCGACTGCTTCC	2100
Qy	2101	GGAAGCACCCGAGGCTACTTACACAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT	2160
Db	2101	GGAAGCACCCGAGGCTACTTACACAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT	2160
Qy	2161	GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTTTCCATCT	2220
Db	2161	GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTTTCCATCT	2220
Qy	2221	TTAAGGTTAGGATGTATGTGGGGGCGTGGAGCACAGGCTCAATGCCGATGCAATTTGGA	2280
Db	2221	TTAAGGTTAGGATGTATGTGGGGGCGTGGAGCACAGGCTCAATGCCGATGCAATTTGGA	2280

Qy	2281	CTCAGGAGGCGCTGTAACTTTGGAGGACAGGGATAGTCAAGACTCAGCCCGCTGCTGC	2340
Db	2281	CTCAGGAGGCGCTGTAACTTTGGAGGACAGGGATAGTCAAGACTCAGCCCGCTGCTGC	2340
Qy	2341	TGTCATAACAGAGTGGCAGATACTGCCTCTGTCTTTACCACTTACCGGCTTTATCCA	2400
Db	2341	TGTCATAACAGAGTGGCAGATACTGCCTCTGTCTTTACCACTTACCGGCTTTATCCA	2400
Qy	2401	CTGGTTTGATCCATCTCCATCAGAAACATCGTGGACGTGCAATACCTGTATAGGTT	2460
Db	2401	CTGGTTTGATCCATCTCCATCAGAAACATCGTGGACGTGCAATACCTGTATAGGTT	2460
Qy	2461	CAGCGTTTGTCTCTTTGGCAATCAATGGAGTACATCTGTGCTTTTCTCTCTCTG	2520
Db	2461	CAGCGTTTGTCTCTTTGGCAATCAATGGAGTACATCTGTGCTTTTCTCTCTCTG	2520
Qy	2521	CAGACGCGCGGTCTGCTCTGCTGTGATGATGCTGTAGTCAGCCAGGCTGAGCCG	2580
Db	2521	CAGACGCGCGGTCTGCTCTGCTGTGATGATGCTGTAGTCAGCCAGGCTGAGCCG	2580
Qy	2581	CTTTAGAGAACTTGTGTGCTCTCAATGCGCGCTCCGTGGCCGAGCGCATGTATTCT	2640
Db	2581	CTTTAGAGAACTTGTGTGCTCTCAATGCGCGCTCCGTGGCCGAGCGCATGTATTCT	2640
Qy	2641	CTTTCTTGTGTTCTTCTGCGCCGCTCGTACATTAAGGGCAGGCTCGCTCTGGGCGG	2700
Db	2641	CTTTCTTGTGTTCTTCTGCGCCGCTCGTACATTAAGGGCAGGCTCGCTCTGGGCGG	2700
Qy	2701	CGTACGCTTTTATGGGTATGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2760
Db	2701	CGTATGCTTTTATGGGTATGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2760
Qy	2761	CTTACGCTTTGGACCGGAGATGGCTGCATCTGTGCGGGGTGCGGTTCTTGTAGTCTG	2820
Db	2761	CTTACGCTTTGGACCGGAGATGGCTGCATCTGTGCGGGGTGCGGTTCTTGTAGTCTG	2820
Qy	2821	TATCTTGTACCTTGTGCACATCTCAAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCT	2880
Db	2821	TATCTTGTACCTTGTGCACATCTCAAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCT	2880
Qy	2881	AATACCTTTATCACAGAGCGGCGGCACATGCAAGTGTGGGTCCTCTCTCTCTCTCT	2940
Db	2881	AATACCTTTATCACAGAGCGGCGGCACATGCAAGTGTGGGTCCTCTCTCTCTCTCT	2940
Qy	2941	GGGAGCGCCGATGCCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3000
Db	2941	GGGAGCGCCGATGCCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3000
Qy	3001	ACATCACCAAACTCTGCTCGCCATCTCGGCCGCTCATGTGTCTCCAGGCTGGCATAA	3060
Db	3001	ACATCACCAAACTCTGCTCGCCATCTCGGCCGCTCATGTGTCTCCAGGCTGGCATAA	3060
Qy	3061	CGAGAGTCCGTACTTGTGCGGCTCAAGGCTCATCTCTCTCTCTCTCTCTCTCTCTCT	3120
Db	3061	CGAGAGTCCGTACTTGTGCGGCTCAAGGCTCATCTCTCTCTCTCTCTCTCTCTCTCT	3120
Qy	3121	AAGTCGCGGGGTCTATTGTCCAAATGTCTTCTATGAAGTGTGGGCGCTGACAGGTA	3180
Db	3121	AAGTCGCGGGGTCTATTGTCCAAATGTCTTCTATGAAGTGTGGGCGCTGACAGGTA	3180
Qy	3181	CGTACGTTTATACCATCTTACCCACTGCGGGACTGGGCCACGCGGGCTTACGAGACC	3240
Db	3181	CGTACGTTTATACCATCTTACCCACTGCGGGACTGGGCCACGCGGGCTTACGAGACC	3240
Qy	3241	TTGCGGTGGCGGTAGACCGCTGTCTTCTCGGCCATGGAGACCAAGGTCTATCACTTGG	3300
Db	3241	TTGCGGTGGCGGTAGACCGCTGTCTTCTCGGCCATGGAGACCAAGGTCTATCACTTGG	3300
Qy	3301	GAGCAGACCGCTGTGCGGTGGGACATCATCTTGGGTCTACCCGTCTCGCCCCGALAGG	3360
Db	3301	GAGCAGACCGCTGTGCGGTGGGACATCATCTTGGGTCTACCCGTCTCGCCCCGALAGG	3360
Qy	3361	GGAAGGAGATATTTTGTGGACCGGCTGTATGTCTCGAAGGCAAGGTTGGCGACTCTTGT	3420

||||| 3361 GGAAGGAGATATTTTGGACCGGCTGATGATCTCGAAGGGCAAGGGTGGGACTCCCTTG 3420
|||||
||||| 3421 CGCCCATCAGCGCTACTCCCAACAAACGCGGGGGTACTTGGTGTGATCATCACTAGCC 3480
|||||
||||| 3421 CGCCCATCAGCGCTACTCCCAACAAACGCGGGGGTACTTGGTGTGATCATCACTAGCC 3480
|||||
||||| 3481 TCACAGGCGGGACAAGAACAGGTGCGAAGGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
|||||
||||| 3481 TCACAGGCGGGACAAGAACAGGTGCGAAGGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
|||||
||||| 3541 AATCTTTCTCGGCACTGTCATCAAACGCGGTGTGCGACTGTCTTACCATGCGCTGGCT 3600
|||||
||||| 3541 AATCTTTCTCGGCACTGTCATCAAACGCGGTGTGCGACTGTCTTACCATGCGCTGGCT 3600
|||||
||||| 3601 CGAAGACCTAGCCGTCCTCAAAAGGTCCAAATCACCNAATGTACCAATGTAGACCTGG 3660
|||||
||||| 3601 CGAAGACCTAGCCGTCCTCAAAAGGTCCAAATCACCNAATGTACCAATGTAGACCTGG 3660
|||||
||||| 3661 ACCTCGTCGGCTGGCAGGCGCCCGCGGGCGGCTCCATGACACCATGCGCTGTGCA 3720
|||||
||||| 3661 ACCTCGTCGGCTGGCAGGCGCCCGCGGGCGGCTCCATGACACCATGCGCTGTGCA 3720
|||||
||||| 3721 GCTCGGACCTTTACTTGGTACAGAGACATGCTGATGTCAATCCGGTGGCGCGGAGGCG 3780
|||||
||||| 3721 GCTCGGACCTTTACTTGGTACAGAGACATGCTGATGTCAATCCGGTGGCGCGGAGGCG 3780
|||||
||||| 3781 ACAGCGGGGAAGTCTACTCTCCCGAGCGCGCTCTCTACCTGAAAGCTCTCGGGTG 3840
|||||
||||| 3781 ACAGCGGGGAAGTCTACTCTCCCGAGCGCGCTCTCTACCTGAAAGCTCTCTCGGGTG 3840
|||||
||||| 3841 GTCCATTGCTTTGCGCTTCGCGGACGCTGCGGGGCTCTCCGGGCTGCTGTGTGCAACC 3900
|||||
||||| 3841 GTCCATTGCTTTGCGCTTCGCGGACGCTGCGGGGCTCTCCGGGCTGCTGTGTGCAACC 3900
|||||
||||| 3901 GGGGGTTCGCGAAGGCGGTGAGCTTCATACCCGTTGAGTCTATGGAACCTACCATGCGGT 3960
|||||
||||| 3901 GGGGGTTCGCGAAGGCGGTGAGCTTCATACCCGTTGAGTCTATGGAACCTACCATGCGGT 3960
|||||
||||| 3961 CTCGGTCTTTCACAGACAACTCAACCCCCCGGGTGTACCGAGACATTCGAAGTGGCAC 4020
|||||
||||| 3961 CTCGGTCTTTCACAGACAACTCAACCCCCCGGGTGTACCGAGACATTCGAAGTGGCAC 4020
|||||
||||| 4021 ATCTGACACCTCTACTGCGAGCGGCAAGACACCAAGTGGCGGTGGTATGAGGCC 4080
|||||
||||| 4021 ATCTGACACCTCTACTGCGAGCGGCAAGACACCAAGTGGCGGTGGTATGAGGCC 4080
|||||
||||| 4081 AAGGGTACAAGGTGCTCGTCTGAAACCCCGTTCGGTTGCCGCCACCTTAGGGTTTGGGGGT 4140
|||||
||||| 4081 AAGGGTACAAGGTGCTCGTCTGAAACCCCGTTCGGTTGCCGCCACCTTAGGGTTTGGGGGT 4140
|||||
||||| 4141 ATATGTCCAAGGACACCGGTATCGACCCCTAAACATCAGAACTGGGGTAAGGACCATTAACA 4200
|||||
||||| 4141 ATATGTCCAAGGACACCGGTATCGACCCCTAAACATCAGAACTGGGGTAAGGACCATTAACA 4200
|||||
||||| 4201 CGGGCGGCTCCATTAGCTACTCCACTATGGCAAGTTCCTTCGCGACGGTGGCTGTCTG 4260
|||||
||||| 4201 CGGGCGGCTCCATTAGCTACTCCACTATGGCAAGTTCCTTCGCGACGGTGGCTGTCTG 4260
|||||
||||| 4261 GGGCGGCTTATGACATCATATAATGTGATGAGTGGCCACTCAACTGACTCGACTACCATCT 4320
|||||
||||| 4261 GGGCGGCTTATGACATCATATAATGTGATGAGTGGCCACTCAACTGACTCGACTACCATCT 4320
|||||
||||| 4321 TGGGCATTCGGCAAGTCTCGGACCAAGCGGAGACGGCTGGAGCGCGGCTGCTGTGCTCG 4380
|||||
||||| 4321 TGGGCATTCGGCAAGTCTCGGACCAAGCGGAGACGGCTGGAGCGCGGCTGCTGTGCTCG 4380
|||||
||||| 4381 CCACCGCTACACCTCCGGGATCGGTTACCGTGGCCACACCCCATATCGAGGAAATAGGCC 4440
|||||
||||| 4381 CCACCGCTACACCTCCGGGATCGGTTACCGTGGCCACACCCCATATATCGAGGAAATAGGCC 4440
|||||
||||| 4441 TGTCCCAACATGAGAGATCCCTTCTATGCGCAAGGCCATCCCTATGAGGCCATCAAGG 4500
|||||

Db 4441 TGTCCAAATGAGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGGCCATCAAGG 4500
QY 4501 GGGGAGGCATCTCATTTTCTGCCATTCCAAGAGAAATGTGACGAGCTCGCCGCAAGC 4560
Db 4501 GGGGAGGCATCTCATTTTCTGCCATTCCAAGAGAAATGTGACGAGCTCGCCGCAAGC 4560
QY 4561 TGACAGGCCTCGGACTGAACCGCTGTAGCATATTAACCGGGGCTTTGATGTGTCGTATAC 4620
Db 4561 TGACAGGCCTCGGACTGAACCGCTGTAGCATATTAACCGGGGCTTTGATGTGTCGTATAC 4620
QY 4621 CGCCTATCGGAGAGCGTCTGTTGTGCTGCAACAGACGCTCTAAATGACGGGTTTCAACGGCG 4680
Db 4621 CGCCTATCGGAGAGCGTCTGTTGTGCTGCAACAGACGCTCTAAATGACGGGTTTCAACGGCG 4680
QY 4681 ATTTTGTAGTCACTGATCGACTGCAATACATGTGTGTCACCGACAGTGTGAGTTCAGCTTGG 4740
Db 4681 ATTTTGTAGTCACTGATCGACTGCAATACATGTGTGTCACCGACAGTGTGAGTTCAGCTTGG 4740
QY 4741 ATCCCACTTTCAACCATTTGAGACGACGACCGCTGCCCCAAGACGCGGTGTGCGCTCGCAAC 4800
Db 4741 ATCCCACTTTCAACCATTTGAGACGACGACCGCTGCCCCAAGACGCGGTGTGCGCTCGCAAC 4800
QY 4801 GCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC 4860
Db 4801 GCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC 4860
QY 4861 GGCCTCGGGCATGTTGATTTCTTCGGTCTGTGAGTGTCTATGACGCGGGTGTGCTT 4920
Db 4861 GGCCTCGGGCATGTTGATTTCTTCGGTCTGTGAGTGTCTATGACGCGGGTGTGCTT 4920
QY 4921 GGTATGAGCTCACGCCGCTGAGACCTCGGTTAGTGTGCGGGCTTACCTAAATACACAG 4980
Db 4921 GGTATGAGCTCACGCCGCTGAGACCTCGGTTAGTGTGCGGGCTTACCTAAATACACAG 4980
QY 4981 GGTTCGCCGTCTGCGAGAACCATCTGGAGTTCTGGGAGCGTCTTTCACAGGCGCTCACCC 5040
Db 4981 GGTTCGCCGTCTGCGAGAACCATCTGGAGTTCTGGGAGCGTCTTTCACAGGCGCTCACCC 5040
QY 5041 ACATAGATGCCACTTCTGTCGCCAGACTAAACAGGAGGAGACAACTTTCTTACCTCTG 5100
Db 5041 ACATAGATGCCACTTCTGTCGCCAGACTAAACAGGAGGAGACAACTTTCTTACCTCTG 5100
QY 5101 TGGCATATCAAGCTACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCATCTGCGGACCCAAA 5160
Db 5101 TGGCATATCAAGCTACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCATCTGCGGACCCAAA 5160
QY 5161 TGTGGAAGTGTCTCATAGGCTGAAACCTACACTGCAAGGCGCCAAACACCCCTGCTGTATA 5220
Db 5161 TGTGGAAGTGTCTCATAGGCTGAAACCTACACTGCAAGGCGCCAAACACCCCTGCTGTATA 5220
QY 5221 GGCTAGGAGCGCTCCAAATGAGGTCTATCCACACACCCCATTAACATAACATCATGG 5280
Db 5221 GGCTAGGAGCGCTCCAAATGAGGTCTATCCACACACCCCATTAACATAACATCATGG 5280
QY 5281 CATGATCTCGGCTGACCTGGAGGTGCTCACTAGCACCTGGGTGTGCTAGCGGAGTCC 5340
Db 5281 CATGATCTCGGCTGACCTGGAGGTGCTCACTAGCACCTGGGTGTGCTAGCGGAGTCC 5340
QY 5341 TTGCAAGCTTTGGCGCATACTGCTGTGACGACGAGGAGTGTGCTGATGAGGAGGATCA 5400
Db 5341 TTGCAAGCTTTGGCGCATACTGCTGTGACGACGAGGAGTGTGCTGATGAGGAGGATCA 5400
QY 5401 TCTTGTCCGGGAAGCCAGCTGCTGTTCCGACAGGAGTCTCTACAGGAGTTCGATG 5460
Db 5401 TCTTGTCCGGGAAGCCAGCTGCTGTTCCGACAGGAGTCTCTACAGGAGTTCGATG 5460
QY 5461 AGATGGAAGTGTGCTCTCAACATTTCTTATCATCGACAGGGAATGCAAGTTCGCGGAGC 5520
Db 5461 AGATGGAAGTGTGCTCTCAACATTTCTTATCATCGACAGGGAATGCAAGTTCGCGGAGC 5520
QY 5521 AATTCAGCAAAAGGCGCTCGGGTTGTCGAAAACGGCCACCAAGCAAGCGGAGGCTGCTG 5580
Db 5521 AATTCAGCAAAAGGCGCTCGGGTTGTCGAAAACGGCCACCAAGCAAGCGGAGGCTGCTG 5580

QY	5581	CTCCGTCGTGGAGTCCAAAGTCGGCAGCCCTTGAGACCTTCTGSGCGAAGACACATGTGGA	5640
DB	5581	CTCCCGTCGTGGAGTCCAAAGTCGGCAGCCCTTGAGACCTTCTGSGCGAAGACACATGTGGA	5640
QY	5641	ATTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTCGAAACCCCGGA	5700
DB	5641	ATTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTCGAAACCCCGGA	5700
QY	5701	TAGCATCATGTATGGCAATTACAGTTCTATCAGTAGCCGCTCACCACCCAAAACACC	5760
DB	5701	TAGCATCATGTATGGCAATTACAGTTCTATCAGTAGCCGCTCACCACCCAAAACACC	5760
QY	5761	TCTGTTTAAACATCTTTGGGGGATGGTGCGTCGCCAACTCGCTCTCTCCAGCGCTGCGT	5820
DB	5761	TCTGTTTAAACATCTTTGGGGGATGGTGCGTGCCCAACTCGCTCTCTCCAGCGCTGCGT	5820
QY	5821	CAGCTTTCGTGGCGCGCGCATCGCCGAGCGGCTGTGGCAGCATPAGCCCTTGGGAAGG	5880
DB	5821	CAGCTTTCGTGGCGCGCGCATCGCCGAGCGGCTGTGGCAGCATPAGCCCTTGGGAAGG	5880
QY	5881	TGCTCTGTGACATCTTTGGCGGGCTATGGGCGCAGGGGTAGCGGGCGCACTCTGTGGCCCTTTA	5940
DB	5881	TGCTCTGTGACATCTTTGGCGGGCTATGGGCGCAGGGGTAGCGGGCGCACTCTGTGGCCCTTTA	5940
QY	5941	AGSTCATGCGCGGAGGTGCCCTCCACGAGAACCTGGTCAACTTACTCCCTGCCATCC	6000
DB	5941	AGSTCATGCGCGGAGGTGCCCTCCACGAGAACCTGGTCAACTTACTCCCTGCCATCC	6000
QY	6001	TCCTCTCTGTGCCCTCGTCTCGGGTCTGTGCGCAGCAATPACTTCGCTCGGCACGTGG	6060
DB	6001	TCCTCTCTGTGCCCTCGTCTCGGGTCTGTGCGCAGCAATPACTTCGCTCGGCACGTGG	6060
QY	6061	GCCCGGAGAGGGGGCTGTGCATGTGATGAACCGGGCTGATPAGGTCCTTCGCGGGGTA	6120
DB	6061	GCCCGGAGAGGGGGCTGTGCATGTGATGAACCGGGCTGATPAGGTCCTTCGCGGGGTA	6120
QY	6121	ACACAGTCTCCCTACGCACTATGTGCTGAGAGCGACGCTGCGACGCGTGCACTCAGA	6180
DB	6121	ACACAGTCTCCCTACGCACTATGTGCTGAGAGCGACGCTGCGACGCGTGCACTCAGA	6180
QY	6181	TCCTCTCTAGCCTTACCATCACTCAACTGCTGAACGGGCTCCACAGTGGATTAATGAGG	6240
DB	6181	TCCTCTCTAGCCTTACCATCACTCAACTGCTGAACGGGCTCCACAGTGGATTAATGAGG	6240
QY	6241	ACTGCTCTAGCCCATGTCCGGCTCGTGCTTAAGGGATGTTTGGATTTGCATATGCACGG	6300
DB	6241	ACTGCTCTAGCCCATGTCCGGCTCGTGCTTAAGGGATGTTTGGATTTGCATATGCACGG	6300
QY	6301	TGTTGACTGACTTCAAGACCTTGGCTCAGTCCAAACTCTGCGCGGTTACCGGAGTCC	6360
DB	6301	TGTTGACTGACTTCAAGACCTTGGCTCAGTCCAAACTCTGCGCGGTTACCGGAGTCC	6360
QY	6361	CTTTCCTGTCATGCCAACGCGGGTACAGGGAGTCTGCGGGGGGACGGCATCATGCAAA	6420
DB	6361	CTTTCCTGTCATGCCAACGCGGGTACAGGGAGTCTGCGGGGGGACGGCATCATGCAAA	6420
QY	6421	CCACTGCCCCATGGGGACACAGATCGCGGACATGTCAAAAACGGTTTCCATGAGGATCG	6480
DB	6421	CCACTGCCCCATGGGGACACAGATCGCGGACATGTCAAAAACGGTTTCCATGAGGATCG	6480
QY	6481	TAGGGCCTTAGAACCTGCGACGAACAGTGGCAACGGTTCCTCCCATCAACGCGATACACCA	6540
DB	6481	TAGGGCCTTAGAACCTGCGACGAACAGTGGCAACGGTTCCTCCCATCAACGCGATACACCA	6540
QY	6541	CGGGACCTTGACACCTTCCCGGGGCCCACTATTCCAGGGCGCTATGGCGGTGGCTG	6600
DB	6541	CGGGACCTTGACACCTTCCCGGGGCCCACTATTCCAGGGCGCTATGGCGGTGGCTG	6600
QY	6601	CTGAGGAGTACGTGAGGTTACGGGTGTGGGGGATTTTCCACTACGTGACGGGCATGACCA	6660
DB	6601	CTGAGGAGTACGTGAGGTTACGGGTGTGGGGGATTTTCCACTACGTGACGGGCATGACCA	6660

Qy	6661	CTGACAAAGTAAAGTGCCCATGCTCCAGTTTCGGGCCCGCCCGAATTTCTTACGGAGGTGGATG	6720
Db	6661	CTGACAAAGTAAAGTGCCCATGCTCCAGTTTCGGGCCCGCCCGAATTTCTTACGGAGGTGGATG	6720
Qy	6721	GAGTGGGTGGACAGGTACGCTCCGGCTGCAAACTCTTCTACGGAGGAGCGTCACT	6780
Db	6721	GAGTGGGTGGACAGGTACGCTCCGGCTGCAAACTCTTCTTACGGAGGAGCGTCACT	6780
Qy	6781	TCCAGGTTCGGGCTCAACAAATACTTTGGTTCGGGTTCGAGCTCCCATGCGAGCCGGAACCGG	6840
Db	6781	TCCAGGTTCGGGCTCAACAAATACTTTGGTTCGGGTTCGAGCTCCCATGCGAGCCGGAACCGG	6840
Qy	6841	ACGTAAACAGTCTTACTTTCCATGCTCACGATCCCTCCACATTTACGACAGAGACGGCTA	6900
Db	6841	ACGTAAACAGTCTTACTTTCCATGCTCACGATCCCTCCACATTTACGACAGAGACGGCTA	6900
Qy	6901	AGCGTAGGCTGGCTAGAGGGTCTCCCGCTTTTACGACGCTCATCAGCTAGCCAGTTGT	6960
Db	6901	AGCGTAGGCTGGCTAGAGGGTCTCCCGCTTTTACGACGCTCATCAGCTAGCCAGTTGT	6960
Qy	6961	CTGCGCTTTCTTTGAAGCGACATGCACTACCAACCATGATCCCCTGGAGCGCTGACTCA	7020
Db	6961	CTGCGCTTTCTTTGAAGCGACATGCACTACCAACCATGATCCCCTGGAGCGCTGACTCA	7020
Qy	7021	TCGAGGCCAACCTCTTTGTGGCGGAGAGATGGCGGAAACATCATCTCGCGTGGAGTCA	7080
Db	7021	TCGAGGCCAACCTCTTTGTGGCGGAGAGATGGCGGAAACATCATCTCGCGTGGAGTCA	7080
Qy	7081	AGATATAGGTAGTAATTTCTGACTCTTTTCCAAACCGCTTTCACGCGAGGGGGATGAGAGGG	7140
Db	7081	AGATATAGGTAGTAATTTCTGACTCTTTTCCAAACCGCTTTCACGCGAGGGGGATGAGAGGG	7140
Qy	7141	AGATATCCGTTCGGCGGAGATCCTGCGAAATTCAGGAAAGTTCCCTCAGCGTTGCGCCA	7200
Db	7141	AGATATCCGTTCGGCGGAGATCCTGCGAAATTCAGGAAAGTTCCCTCAGCGTTGCGCCA	7200
Qy	7201	TATGGGACGCGCGGACTCAATCTCTCCTAGAGTCTTGGAGGACCCGGAGTAC	7260
Db	7201	TATGGGACGCGCGGACTCAATCTCTCCTAGAGTCTTGGAGGACCCGGAGTAC	7260
Qy	7261	TCCCTTCGGTGGTACGGATGCCATTGCCACTTACCAAGGCTCTTCCAAATACCACTC	7320
Db	7261	TCCCTTCGGTGGTACGGATGCCATTGCCACTTACCAAGGCTCTTCCAAATACCACTC	7320
Qy	7321	CACGGAGAAAGGACGGTTGCTCTGACAGAAATCCAAATGTCTTTCCTTGGCGGAGC	7380
Db	7321	CACGGAGAAAGGACGGTTGCTCTGACAGAAATCCAAATGTCTTTCCTTGGCGGAGC	7380
Qy	7381	TGCGCACTAAGACCTTCGGTAGCTCCGGATCGTCGGCGTTGATAGCGGCACCGCGACCG	7440
Db	7381	TGCGCACTAAGACCTTCGGTAGCTCCGGATCGTCGGCGTTGATAGCGGCACCGCGACCG	7440
Qy	7441	CCCTTCCTGACTCGCTCCCGACGAGTGAAGGATCCGAGCTTGAGCTGCTACTCCT	7500
Db	7441	CCCTTCCTGACTCGCTCCCGACGAGTGAAGGATCCGAGCTTGAGCTGCTACTCCT	7500
Qy	7501	CCATGCCCTTTGAAGGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTA	7560
Db	7501	CCATGCCCTTTGAAGGGGAGCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTA	7560
Qy	7561	CCGTGAGTGAGAGCTTAGTGAGGATGTCGTCTGCTCTTAATGCTCTATACGTGGACAG	7620
Db	7561	CCGTGAGTGAGAGCTTAGTGAGGATGTCGTCTGCTCTTAATGCTCTATACGTGGACAG	7620
Qy	7621	GGCGCTGATCACGCCATGCCCTCGGAGGAAAGTACGTCCCATCAACCCGTTGAGCA	7680
Db	7621	GGCGCTGATCACGCCATGCCCTCGGAGGAAAGTACGTCCCATCAACCCGTTGAGCA	7680
Qy	7681	ACTCTTTGCTGGTCAACAACATGGTCTACGGCACCAACATCCCGCAGCGCAAGCTCC	7740
Db	7681	ACTCTTTGCTGGTCAACAACATGGTCTACGGCACCAACATCCCGCAGCGCAAGCTCC	7740
Qy	7741	GGCAGAAAGGTACCTTTTGACAGATTTGCAAGTCTCTGGATGATCATTTACGGGACGTAC	7800

FEATURES	7 Center Dr., Building 7, Room 201, Bethesda, MD 20892, USA			
	Location/Qualifiers			
	1. .9595			
	/organism="Hepatitis C virus"			
	/mol_type="genomic RNA"			
CDS	/strain="HC-J4"			
	/isolate="pcv-J4L68"			
	/db_xref="taxon:11103"			
	/notes="infectious full-length cDNA"			
	342. .9374			
ORIGIN	/codon_start=1			
	/product="polyprotein"			
	/protein_id="AAC15722.1"			
	/db_xref="GI:3098633"			
	/translation="MSTNPKPQRKTRNTNRPDQVFKPGGQIVGGVYLLPRRGRPL GVNRKTPRSRORRROPVKPAREPGRWAQPCYPMPLYCNEGLGAGWAGLLSPRG SPSNGKTPRSRNLKVIDTLCTGCFADLMGYILVPGVAPLGGAARALAGHVRVLG GVNATGNLPGCSFSFLALLSLCLTIPASAEVENVSGIYHVTNDCSNSSIVYEAD VNHITPCVPCVQEGNSRCWALTPILARNASVPTTIRRHVDLLVGTAAFCASMY VGLDCSILFVLSQFTFSPRRHETQDCNSCTPGVSHGRHMAWMNNSPTTLLR SOLLRIPOAVVMAGAHGVLAGLAYSMGVNKAIVLLVALLFAGVDETHTTGRVA GHTTSFTLSFSSGASQKLTQVNTGSMHNRNTALNCNDSLOTQGFPAALFYAHKNS GCPERASCRPIDWFAQWGPIYTKPNSDORPYCMHYAPRCGVVPASOVCVGPVC FTSPVGGTIDRSQVPTYSBNETDYMLLNTRPPQGNWFGCTWMASTGFTKTCGG PCNIGGVNRTLICPTCFRKHPEATYTKCGSGPMLTPRCLVDVPIRHWHPYLTCLNF SIFKVMQSGEHRINACNWTGRERCNLEDRSSELSPLLSTLSTTEWLPICAFTHL PALSTGLIHQNIVDQVLYGVGSFAVFAIKWEYLILLFLLLDARVACALWMLL IAQEALENLNLVNASVAGAHILSFVFFCAAWYIKGRLAGPAAGVACVGFWPLL LLALPPRAYALDREMAASCGGANVLGLVFLILSPYKYVFLRLIWMLOYFIFTRAEB MQWVPPNVNRRGRDAILILTCVAHPLELIDTLGLLAILGLPMLQAGITKVPVFR AOGILRACMLRVAGGVYQVFMKLGALTGYVYHNLTPLRDMAHAGLRDLAVAVE PVVFSAMETKVIWIGADTRACGDIILGLPVARSARKEIFLGPADSLEGQMRLLAIP ASYQOTRGVLGCIITSLTRGDNQVEGVVSTATQSPLATCINGCVMTVTVHGASK TLAGKGPITQWYTNLDLVGQAAPGARSMTPCSCSSDLYLVTRHADVIVPVRRG DSRGLSPRPVYLKSGSGPLLCFSGHVGVFRAAVCTRGVAKAVDFIPVESMETT MRGPFVNTSPFAVPTFOVHLHPTGSGKTKVPAAYAAQGYKVLVNPESVAATL GFGAYNSKSTIGADRTGTGVTITGGSITYTYGKLADGGCGGAYDIIICDECH TDSITLIGLTQAEHAKKARLWILATATPGSVTPHPNTEIIGLSNNGIPFYGK AIPLEATKGRHLIFCHSKACDELAAKLTGLGLNAVAYYRGLDVSVIPIIDGVVVA TDALMTGTFDPSVDICNTCVTQTVDFSLDPTFTIETTVPQDAVSRORRGTRCG RSGIYRVTGPERPSGDFSSVJCECDAGCAWYELTPAETSVRLRAYLNTFGLPVCQ DHLEFVESVFTGLTHIDAHFLTSQTKAGDNFPYLVAQVATCARAQAPPPSPDMWK LIRLPLTGHPTPLVRLGAVQNEVILITKYINMACSADLEVTVTWLVGGVLA ALAACTLTGSPVIVGRIILSGKAVVUPREILYQFDEMEECASQLPYIEQGMOLAE QFKQKALGILLQATKQAEAPAVPESKWRALETFMKHMNRFISGIFYLAGLSTLPGN PATASLWATASITSPSTQNTQLLENILGCVAAQLAPPASASAFVAGIAGNAGVSI GLGVVLDIAGAGVAGVALAFKPVMSGEVSTEDLVNLLPAILSPAGVGVVCAVNA ILRRHVGEGGAVQMMNRLJAFASRGHNVSPTHYPVESDAARVTOILSLTILQLLK RLHWINEDCTSPKCSWLRDWDVICTVLDTFKTLQSKLLPLRGVPLFSCQRYK GVNRDGI MOTTCPGAQIAHGVKNSMRIVGPTICNTWHGTFPINAVITGCTPSP APNYSRALNRVAEEYVTVRGDPIYVTGTTNDVKCPCOVPAPEFTEVDGVRILHR YAPACKLLREDVTVGLNQYLVGSPCEPDPVTLTSLMTDPSHITAETAKRRL ARGSPPLSSASSQSASPLKSLITDTHDPSDADLIEANLLWRQMGNIITRVESEN KVLVLDSEFLHAGDEBRESVAEILRKSRKFPSPALPTIWARPDNPPLLEKWDPDY VPVWPCSLPTKAPPIPPRKRTVLTVESVSSALABELAKTKTSGSSGSAVDSGT ATAPLDIASDGDGSDVESYSSMPLEGGPDPDLSGSMSTVSEAEEDVVCSSMS YTWGALITPCAAEESKLPINPLNSLLRHHNMVYATTSRSASLRKQKVTDFORLVLD DHYRDLKEMKAKATVKAULLSIEEACKLTPPHSAKSKFGYGAKOVRLNLSRNVHI RSVMDELEDDETETIDTIIWAKSVCPQEKGRKPARLIVFDPLGVRCVCKMALYD VWSTLPQAVMSGFGFOSYKQRFVFLVNTWKSCKPMGFSYDTRCFDSTVTSRIRV EESIYQCCDLAPEARQAIRSLTERLYIGGLPTNSKQNGCYRRACRAGSVLTSCGNL TCYLKATPAACPAAKLOCTWMLVNGDLVILCESAGTQEDAAALRAETAMTRYSAPEG DPPOFEYDELELITSCSNVSVAHDASKRYVITLTDPTTIPARAANETARHTPINSLH GNIIMYAPTWARMLMTHFSSILLAQEYKLDCCQIYGACYSIFELDLPQIIIEHLH FLNAVTLHSSYSPGIBNVASCLRLKGVPLPRTWRHRSVRKALLSQGGRAAATCGRYL FNMAVTKLKLTPIPAAASQLDLSGWFVAGVSGGDIYHLSLRARPRWPFPLCLLLLSVGV GIYLLPNR"			
	Query Match	100.0%	Score 9595;	DB 13; Length 9595;
	Best Local Similarity	100.0%;	Prod. No. 0;	
	Matches 9595; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;

Db |||||CGCCAGGAATGCCAGCGTCCCACTAGCAACAATACGACGCCACGCTCGACTTGTCTCGTTG 1140
Qy |||||GGACGGCTGCTTTCTGCTCCGCTAGTACGTGGGGATCTCTCGGATCTATTTTCCCTCG 1200
Db |||||GGACGGCTGCTTTCTGCTCCGCTAGTACGTGGGGATCTCTCGGATCTATTTTCCCTCG 1200
Qy |||||TCTCCAGCTGTTCACTTTCTCGCTCGCCGCTGAGACAGTGCAGGACTGCAACTGCT 1260
Db |||||TCTCCAGCTGTTCACTTTCTCGCTCGCCGCTGAGACAGTGCAGGACTGCAACTGCT 1260
Qy |||||CAATCTATCCCGGCGCATGTATCAGGTCAACGATGGCTTGGGATATGATGATGAACCTGGT 1320
Db |||||CAATCTATCCCGGCGCATGTATCAGGTCAACGATGGCTTGGGATATGATGATGAACCTGGT 1320
Qy |||||CACCTACAAACAGCCCTAGTGGTGCAGAGTTCGCGATCCCAAGCTGTCGTTGGACA 1380
Db |||||CACCTACAAACAGCCCTAGTGGTGCAGAGTTCGCGATCCCAAGCTGTCGTTGGACA 1380
Qy |||||TGGTGGCGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGCTTACTATTTCCATGGTAGGGA 1440
Db |||||TGGTGGCGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGCTTACTATTTCCATGGTAGGGA 1440
Qy |||||ACTGGGCTAAGGTTCTGATTTGGGCTACTCTTTGGCGGCTTCAACGAGAGACCCACA 1500
Db |||||ACTGGGCTAAGGTTCTGATTTGGGCTACTCTTTGGCGGCTTCAACGAGAGACCCACA 1500
Qy |||||CGACGGGAGGTTGGCGGCGCACACACTCGGGTTCACGTCCTTTCTCATCTGGGG 1560
Db |||||CGACGGGAGGTTGGCGGCGCACACACTCGGGTTCACGTCCTTTCTCATCTGGGG 1560
Qy |||||CGTCTCAGAAAATCCAGCTTTGTAATACCAACGGCAGCTGGCACATCAACAGGACTGCC 1620
Db |||||CGTCTCAGAAAATCCAGCTTTGTAATACCAACGGCAGCTGGCACATCAACAGGACTGCC 1620
Qy |||||TAAATTTGAATGACTCCCTCCAAATCGGGTCTTTTGGCGGCTGTTTACGACACAAGT 1680
Db |||||TAAATTTGAATGACTCCCTCCAAATCGGGTCTTTTGGCGGCTGTTTACGACACAAGT 1680
Qy |||||TCAACTCGTCCGGTGGCCGGAGCGCATGGCCAGCTGGCCGCCATGACTGGTTCGCCC 1740
Db |||||TCAACTCGTCCGGTGGCCGGAGCGCATGGCCAGCTGGCCGCCATGACTGGTTCGCCC 1740
Qy |||||AGGGTGGGGCCCCATCACTATCTAAGCCTAAACAGCTCGGATCAGAGGCTTATTTGCT 1800
Db |||||AGGGTGGGGCCCCATCACTATCTAAGCCTAAACAGCTCGGATCAGAGGCTTATTTGCT 1800
Qy |||||GGCATTTACCGCTCGACCGTGTGTGTCGTACCCGCTCGCAGGTGTGTGTTCCAGTGT 1860
Db |||||GGCATTTACCGCTCGACCGTGTGTGTCGTACCCGCTCGCAGGTGTGTGTTCCAGTGT 1860
Qy |||||ATTGTTTACCCCAAGCCCTGTTGGTGGGACCAACCGATCGTTCCGTTGCCCTACGT 1920
Db |||||ATTGTTTACCCCAAGCCCTGTTGGTGGGACCAACCGATCGTTCCGTTGCCCTACGT 1920
Qy |||||ATAGCTGGGGGAGATGAGACAGCTGATGCTCTCTCAACAAACAGCTCGCCGACCAAG 1980
Db |||||ATAGCTGGGGGAGATGAGACAGCTGATGCTCTCTCAACAAACAGCTCGCCGACCAAG 1980
Qy |||||GCAACTGGTTCCGCTGTATCATGGATGAATAGTACTGGGTTCACTAAGACGTCGGAGGTC 2040
Db |||||GCAACTGGTTCCGCTGTATCATGGATGAATAGTACTGGGTTCACTAAGACGTCGGAGGTC 2040
Qy |||||CCCCGTGTAACATCGGGGGGGTGGTAAACCGACCTTGTATCTGCCCGACCGACTGCTTCC 2100
Db |||||CCCCGTGTAACATCGGGGGGGTGGTAAACCGACCTTGTATCTGCCCGACCGACTGCTTCC 2100
Qy |||||GGAAGCACCCGAGGCTACTTACAAAAATGTGGCTCGGGGCCCTGGTTGACACTAGGT 2160
Db |||||GGAAGCACCCGAGGCTACTTACAAAAATGTGGCTCGGGGCCCTGGTTGACACTAGGT 2160
Qy |||||GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTCCATCT 2220

Db |||||GCCTAGTAGACTACCCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTCCATCT 2260
Qy |||||TTAAGGTTAGGATGATGTGGGGGGCTGGAGACAGGCTCAATGCGCGCATGCAATTTGGA 2280
Db |||||TTAAGGTTAGGATGATGTGGGGGGCTGGAGACAGGCTCAATGCGCGCATGCAATTTGGA 2280
Qy |||||CTCAGGAGAGCGCTGTAACTTTGAGACAGGATAGGTCAAGAACTCAGCCCGCTGCTGC 2340
Db |||||CTCAGGAGAGCGCTGTAACTTTGAGACAGGATAGGTCAAGAACTCAGCCCGCTGCTGC 2340
Qy |||||TGTCTCAACAGAGTGGCAGATAGTCCCTGTGCTTTCCACCACCTACCGGCTTTATCCA 2400
Db |||||TGTCTCAACAGAGTGGCAGATAGTCCCTGTGCTTTCCACCACCTACCGGCTTTATCCA 2400
Qy |||||CTGGTTTGTATCCATCTCCATCAGAACTCGTGGACGTGCAATCCTGTACCGGTGTAGGT 2460
Db |||||CTGGTTTGTATCCATCTCCATCAGAACTCGTGGACGTGCAATCCTGTACCGGTGTAGGT 2460
Qy |||||CAGCGTTTGTCTCTTTGCAATGAGGAGTACATCTCTTGTGCTTTTCTCTCTCTCTGG 2520
Db |||||CAGCGTTTGTCTCTTTGCAATGAGGAGTACATCTCTTGTGCTTTTCTCTCTCTCTGG 2520
Qy |||||CAGACGCGCGGTGTGCTGCTCAATGCGGCTCGGTGGATGATGCTGCTATAGCCAGGCTGAGGCGG 2580
Db |||||CAGACGCGCGGTGTGCTGCTCAATGCGGCTCGGTGGATGATGCTGCTATAGCCAGGCTGAGGCGG 2580
Qy |||||CCTTTAGAGAACTTGTGCTGCTCAATGCGGCTCGGTGGAGTACATCTCTTCTCTCTCTCT 2640
Db |||||CCTTTAGAGAACTTGTGCTGCTCAATGCGGCTCGGTGGAGTACATCTCTTCTCTCTCTCT 2640
Qy |||||CCTTTCTTGTGCTTCTTGTGCGCCCTGGTACATTTAAGGGCAGGCTGCTCTCTGGGCGG 2700
Db |||||CCTTTCTTGTGCTTCTTGTGCGCCCTGGTACATTTAAGGGCAGGCTGCTCTCTGGGCGG 2700
Qy |||||CGTATGCTTTTATGGCGTATGGCGGCTGCTCTGCTCCTACTGGCGTTACCAACACAG 2760
Db |||||CGTATGCTTTTATGGCGTATGGCGGCTGCTCTGCTCCTACTGGCGTTACCAACACAG 2760
Qy |||||CTTACGCTTGGACCGGAGATGCTGCTGCGGGGGTGGGTTCTTGTAGTCTCGG 2820
Db |||||CTTACGCTTGGACCGGAGATGCTGCTGCGGGGGTGGGTTCTTGTAGTCTCGG 2820
Qy |||||TATTCTTTGACCTTGTCACTACTACAAAGTGTCTTCTCACTAGGCTCATATGGTGTAC 2880
Db |||||TATTCTTTGACCTTGTCACTACTACAAAGTGTCTTCTCACTAGGCTCATATGGTGTAC 2880
Qy |||||AATACTTTATCACAGAGCGAGGCGACATGCAAGTGTGGTCCCCCCCCCTCAACGTT 2940
Db |||||AATACTTTATCACAGAGCGAGGCGACATGCAAGTGTGGTCCCCCCCCCTCAACGTT 2940
Qy |||||GGGAGGCGCGGATGCCATCATCTCTCACTGCGGTTTCATCCAGAGTTAATTTTG 3000
Db |||||GGGAGGCGCGGATGCCATCATCTCTCACTGCGGTTTCATCCAGAGTTAATTTTG 3000
Qy |||||ACATCAACAAACTCTGCTCGCCATCTCGGCCGCTCATGCTGTCCAGGCTGCATAA 3060
Db |||||ACATCAACAAACTCTGCTCGCCATCTCGGCCGCTCATGCTGTCCAGGCTGCATAA 3060
Qy |||||CGAGAGTCCGCTATCTGCTGCGCGCTCAAGGCTCAITTCGTGTCATGTTAGTGCNA 3120
Db |||||CGAGAGTCCGCTATCTGCTGCGCGCTCAAGGCTCAITTCGTGTCATGTTAGTGCNA 3120
Qy |||||AAGTCGCGGGGTCATTATGTCAAATGTCTTCATGAAGTGGCGGCTGACAGGTA 3180
Db |||||AAGTCGCGGGGTCATTATGTCAAATGTCTTCATGAAGTGGCGGCTGACAGGTA 3180
Qy |||||CGTACGTTTATAACCATCTTACCACCTCGCGGACTGGGCCCAACGCGGCTTACGAGAC 3240
Db |||||CGTACGTTTATAACCATCTTACCACCTCGCGGACTGGGCCCAACGCGGCTTACGAGAC 3240
Qy |||||TTGCGGTGGCGGTAGAGCCCGTCTCTCGGCCATGGAGCAAGGTATCATCCTGGG 3300
Db |||||TTGCGGTGGCGGTAGAGCCCGTCTCTCGGCCATGGAGCAAGGTATCATCCTGGG 3300

QY	3301	GAGCAGACACCGTGGGTGGGACATCATCTTGGGTCTACCCGCTCCGCCGGAAGGG	3360
DB			
QY	3301	GAGCAGACACCGTGGGTGGGACATCATCTTGGGTCTACCCGCTCCGCCGGAAGGG	3360
DB			
QY	3361	GGAGGAGATATTTTGGGACCGGTGATAGTCTCGAAGGGCAAGGGTGGCACTCTTG	3420
DB			
QY	3361	GGAGGAGATATTTTGGGACCGGTGATAGTCTCGAAGGGCAAGGGTGGCACTCTTG	3420
DB			
QY	3421	CGCCCATCAGCGCTACTCCCAACAAACGCGGGCGGTACTTGGTGTGCATCATCACTAGCC	3480
DB			
QY	3421	CGCCCATCAGCGCTACTCCCAACAAACGCGGGCGGTACTTGGTGTGCATCATCACTAGCC	3480
DB			
QY	3481	TCACAGCCGGGCAAGAACCAAGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC	3540
DB			
QY	3481	TCACAGCCGGGCAAGAACCAAGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC	3540
DB			
QY	3541	AATCTTCTGGGACCTGCATCAACGGCGTGTCTGGACTGTCTACCATGGCGCTGGCT	3600
DB			
QY	3541	AATCTTCTGGGACCTGCATCAACGGCGTGTCTGGACTGTCTACCATGGCGCTGGCT	3600
DB			
QY	3601	CGAAGCCCTAGCCGTCCTCAAAAGGTCCTCAATCACCCAAATGTACACCAATGTAGACTGG	3660
DB			
QY	3601	CGAAGCCCTAGCCGTCCTCAAAAGGTCCTCAATCACCCAAATGTACACCAATGTAGACTGG	3660
DB			
QY	3661	ACCTCGTGGGTGGCAGCGCCCCCGGGCGGCTCCATGACACCATGCAGCTGTGGCA	3720
DB			
QY	3721	GCTCGGACCTTACTTGGTCAAGAGACATGCTGATGTCAATTCGGGTGGCGGGCGGCG	3780
DB			
QY	3781	ACAGCGGGAGGTCCTCTCCCCAGGGCCGCTCTACTCAAAAGGTCCTCGGGTG	3840
DB			
QY	3781	ACAGCGGGAGGTCCTCTCCCCAGGGCCGCTCTACTCAAAAGGTCCTCGGGTG	3840
DB			
QY	3841	GTCCATTGCTTGGCCCTTGGGGCAGCGTGTGGGCGTCTTCCGGGCTGCTGTGCAACC	3900
DB			
QY	3841	GTCCATTGCTTGGCCCTTGGGGCAGCGTGTGGGCGTCTTCCGGGCTGCTGTGCAACC	3900
DB			
QY	3901	GGGGGTGCGAAGCGGTGACCTTACATCCGTTGAGTCTATGGAACTACCATGCGGT	3960
DB			
QY	3901	GGGGGTGCGAAGCGGTGACCTTACATCCGTTGAGTCTATGGAACTACCATGCGGT	3960
DB			
QY	3961	CTCGGCTTTCACAGACAACTCAACCCCGCGGTGTACCCGACACATTCCAAGTGGCAC	4020
DB			
QY	3961	CTCGGCTTTCACAGACAACTCAACCCCGCGGTGTACCCGACACATTCCAAGTGGCAC	4020
DB			
QY	4021	ATCTGACGCTCTACTGGCAGCGGCAAGAGCACCAAGTGCCTGCTGATGCAAGCC	4080
DB			
QY	4021	ATCTGACGCTCTACTGGCAGCGGCAAGAGCACCAAGTGCCTGCTGATGCAAGCC	4080
DB			
QY	4081	AAGGGTCAAGGTGCTCGTCTCTGAACCCGTTGGCGCAACCTTAGGGTTTGGGGGT	4140
DB			
QY	4081	AAGGGTCAAGGTGCTCGTCTCTGAACCCGTTGGCGCAACCTTAGGGTTTGGGGGT	4140
DB			
QY	4141	ATATGTCGAAGGCAACCGGTATCGACCTTAACATCAGAACTGGGGTAAGGACCATACCA	4200
DB			
QY	4141	ATATGTCGAAGGCAACCGGTATCGACCTTAACATCAGAACTGGGGTAAGGACCATACCA	4200
DB			
QY	4201	CGGGGGCTCCATTACGTACTCCACTATGGCAAGTTCCTTCCGCAAGCGTGGCTTCTG	4260
DB			
QY	4201	CGGGGGCTCCATTACGTACTCCACTATGGCAAGTTCCTTCCGCAAGCGTGGCTTCTG	4260
DB			
QY	4261	GGGGGGCTTATGACATCAATATGATGATGAGTGCCTCAACTGACTCGACTACCATCT	4320
DB			
QY	4261	GGGGGGCTTATGACATCAATATGATGATGAGTGCCTCAACTGACTCGACTACCATCT	4320
DB			
QY	4321	TGGGATCGGCAACAGTCTCTGGAACCAAGGGGAGACGGCTGGAGCGGCTCGTCTGCTCG	4380
DB			
QY	4321	TGGGATCGGCAACAGTCTCTGGAACCAAGGGGAGACGGCTGGAGCGGCTCGTCTGCTCG	4380
DB			

QY	4381	CCACCGCTACCTCCGGATCGGTTACCGTGCACACCCCAATATCGAGGAAATAGGCC	4440
DB			
QY	4381	CCACCGCTACCTCCGGATCGGTTACCGTGCACACCCCAATATCGAGGAAATAGGCC	4440
DB			
QY	4441	TGTCCAAATGGAGAGATCCCTTCTATGGCAAGCCATCCCAATGGAGGCATCAAG	4500
DB			
QY	4441	TGTCCAAATGGAGAGATCCCTTCTATGGCAAGCCATCCCAATGGAGGCATCAAG	4500
DB			
QY	4501	GGGGAGGATCTCATTTTCTGCCATTCCAAGAAATGTGACGAGCTCGCCGCAAGC	4560
DB			
QY	4501	GGGGAGGATCTCATTTTCTGCCATTCCAAGAAATGTGACGAGCTCGCCGCAAGC	4560
DB			
QY	4561	TGACAGCCTCGAGCTGAACGCTGTAGCATATTACCGGGCCCTTGATGTGCTCATAC	4620
DB			
QY	4561	TGACAGCCTCGAGCTGAACGCTGTAGCATATTACCGGGCCCTTGATGTGCTCATAC	4620
DB			
QY	4621	CGCTATCGAGAGCTGTTTGTCTGGCAACAGACGCTCTAATGACGGGTTTCAACGGCG	4680
DB			
QY	4621	CGCTATCGAGAGCTGTTTGTCTGGCAACAGACGCTCTAATGACGGGTTTCAACGGCG	4680
DB			
QY	4681	ATTTGACTCAGTGATCGACTGCAATACATGTGTACCCAGACAGTCGACTTCAGCTTGG	4740
DB			
QY	4681	ATTTGACTCAGTGATCGACTGCAATACATGTGTACCCAGACAGTCGACTTCAGCTTGG	4740
DB			
QY	4741	ATCCACCTTACCATTTGAGACGACCGCTGCCCAAGACGCGTCTCGCGTTCGCAAC	4800
DB			
QY	4741	ATCCACCTTACCATTTGAGACGACCGCTGCCCAAGACGCGTCTCGCGTTCGCAAC	4800
DB			
QY	4801	GGCAGGTAGAACTGGCAGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC	4860
DB			
QY	4801	GGCAGGTAGAACTGGCAGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC	4860
DB			
QY	4861	GGCCTCGGGCATTTTCGATTTCGTTCTGTGTGAGTGTCTATGACGCGGGCTGTGCTT	4920
DB			
QY	4861	GGCCTCGGGCATTTTCGATTTCGTTCTGTGTGAGTGTCTATGACGCGGGCTGTGCTT	4920
DB			
QY	4921	GGTATGAGCTCAGCCCGCTGAGACCTCGGTAGTTTGGGGTCTTAAATACACAG	4980
DB			
QY	4921	GGTATGAGCTCAGCCCGCTGAGACCTCGGTAGTTTGGGGTCTTAAATACACAG	4980
DB			
QY	4981	GGTTCGCGCTCTGCCAGGACCACTTGGAGTCTTGGGAGAGCGTCTTACAGGCTCACCC	5040
DB			
QY	4981	GGTTCGCGCTCTGCCAGGACCACTTGGAGTCTTGGGAGAGCGTCTTACAGGCTCACCC	5040
DB			
QY	5041	ACATAGATGCCACTTCTCTGCCAGACTAAACAGGAGAGAGACAACTTTCCTTACCTGG	5100
DB			
QY	5041	ACATAGATGCCACTTCTCTGCCAGACTAAACAGGAGAGAGACAACTTTCCTTACCTGG	5100
DB			
QY	5101	TGGCATATCAAGCTACAGTGTGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAA	5160
DB			
QY	5101	TGGCATATCAAGCTACAGTGTGCCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAAA	5160
DB			
QY	5161	TGTGGAAGTGTCTCATACGCTGAAACCTACACTGACGCGGCCAACACCCCTGTGTATA	5220
DB			
QY	5161	TGTGGAAGTGTCTCATACGCTGAAACCTACACTGACGCGGCCAACACCCCTGTGTATA	5220
DB			
QY	5221	GGCTAGGAGCGCTCCAAAATGAGGTCTCTTCAACACCCCAATAACTAAATACATATGG	5280
DB			
QY	5221	GGCTAGGAGCGCTCCAAAATGAGGTCTCTTCAACACCCCAATAACTAAATACATATGG	5280
DB			
QY	5281	CATGCATGTGGCTGACCTGGAGTGTCTACTAGACCTGGGTCTGGTAGCGGAGTCC	5340
DB			
QY	5281	CATGCATGTGGCTGACCTGGAGTGTCTACTAGACCTGGGTCTGGTAGCGGAGTCC	5340
DB			
QY	5341	TTGAGCTTTGGCGCATACTGCTGACGACAGCAGTGTGTCTATTGTGGGAGGATCA	5400
DB			
QY	5341	TTGAGCTTTGGCGCATACTGCTGACGACAGCAGTGTGTCTATTGTGGGAGGATCA	5400
DB			
QY	5401	TCTTGTCCGGAAGCGAGCTGTCTTCCGAGAGGAGTCTCTTACAGAGTTCGATG	5460
DB			
QY	5401	TCTTGTCCGGAAGCGAGCTGTCTTCCGAGAGGAGTCTCTTACAGAGTTCGATG	5460
DB			
QY	5461	AGATGGAAGAGTGTGCTCTCAAACTTCTTATACGACGAGGAGATGAGCTCGCCGAGC	5520
DB			

Db 5461 AGATGGAAGAGTGTGCCTCACAACTTCTCTTATCATCGAGCAGGGAATGCAGCTGCGCGAGC 5520
QY 5521 AATTCAAGCAAGAGCGCTCGGTTGTTGCAAAACGGCCACCAAGCAAGCGGAGGCTGCTG 5580
Db 5521 AATTCAAGCAAGAGCGCTCGGTTGTTGCAAAACGGCCACCAAGCAAGCGGAGGCTGCTG 5580
QY 5581 CTCCTCGTGGTGAAGTCCAAAGTGGCGAGCCCTTGAGACCTTCTGGCGGAAGACACATGTTGGA 5640
Db 5581 CTCCTCGTGGTGAAGTCCAAAGTGGCGAGCCCTTGAGACCTTCTGGCGGAAGACACATGTTGGA 5640
QY 5641 APTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCAATCTGCTGCTGGAACCCCGGGA 5700
Db 5641 APTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCAATCTGCTGCTGGAACCCCGGGA 5700
QY 5701 TAGCATCATTGATGCAATACAGCTTCTATCACTAGCCCGCTACACACCCCAACACCC 5760
Db 5701 TAGCATCATTGATGCAATACAGCTTCTATCACTAGCCCGCTACACACCCCAACACCC 5760
QY 5761 TCCTGTTTAAATCTTGGGGGATGGGTGCTGCCAATCGCTCCTCCAGCGCTGGT 5820
Db 5761 TCCTGTTTAAATCTTGGGGGATGGGTGCTGCCAATCGCTCCTCCAGCGCTGGT 5820
QY 5821 CAGCTTTTCTGGGCGCCGAGCATCGCCGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAG 5880
Db 5821 CAGCTTTTCTGGGCGCCGAGCATCGCCGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAG 5880
QY 5881 TGCTCGTGACATCTTGGGGGCTATGGGAGGGGTAGCCGCGCACTCGTGGGCTTTA 5940
Db 5881 TGCTCGTGACATCTTGGGGGCTATGGGAGGGGTAGCCGCGCACTCGTGGGCTTTA 5940
QY 5941 AGGTATGAGCGCGAGGTGCCCTCCACCGAGACCTGCTCAACTTACTTCCCTGCCATCC 6000
Db 5941 AGGTATGAGCGCGAGGTGCCCTCCACCGAGACCTGCTCAACTTACTTCCCTGCCATCC 6000
QY 6001 TCTCTCTGCTGCTGCTGCTGCGGGTCTGTGCGGAGCAATACTTGGCTCGGCACTGG 6060
Db 6001 TCTCTCTGCTGCTGCTGCTGCGGGTCTGTGCGGAGCAATACTTGGCTCGGCACTGG 6060
QY 6061 GCCCGGAGAGGGGCTGTGCAAGTGAATGAACCGGCTGATAGCGTTTGGCGGGGTA 6120
Db 6061 GCCCGGAGAGGGGCTGTGCAAGTGAATGAACCGGCTGATAGCGTTTGGCGGGGTA 6120
QY 6121 ACCAGTCTCCCTACGCACTATGCTGAGAGGAGCGCTCAGCAGGTGCACTCAGA 6180
Db 6121 ACCAGTCTCCCTACGCACTATGCTGAGAGGAGCGCTCAGCAGGTGCACTCAGA 6180
QY 6181 TCCTCTAGCCCTTACCATCACTCAACTGCTGAAGCGGCTCCACCAGTGGATTAATGAGG 6240
Db 6181 TCCTCTAGCCCTTACCATCACTCAACTGCTGAAGCGGCTCCACCAGTGGATTAATGAGG 6240
QY 6241 ACTGCTCTAGCGCATGCTCGGGTCTGTGCTAAGGATGTTTGGATTTGGATATGACGG 6300
Db 6241 ACTGCTCTAGCGCATGCTCGGGTCTGTGCTAAGGATGTTTGGATTTGGATATGACGG 6300
QY 6301 TGTGACTCACTTCAAGACTGCTGCTCCAGTCCAACTCTGCGCGGTTTACCGGAGTCC 6360
Db 6301 TGTGACTCACTTCAAGACTGCTGCTCCAGTCCAACTCTGCGCGGTTTACCGGAGTCC 6360
QY 6361 CTTTTCCTGTCATGCCAACCGGGGTACAAGGAGTCTGGCGGGGACCGGCATCATGCAAA 6420
Db 6361 CTTTTCCTGTCATGCCAACCGGGGTACAAGGAGTCTGGCGGGGACCGGCATCATGCAAA 6420
QY 6421 CCACCTGCCCATGCGGAGACAGATCGCGGACATGTCAAAAACGTTTCCATGAGATCG 6480
Db 6421 CCACCTGCCCATGCGGAGACAGATCGCGGACATGTCAAAAACGTTTCCATGAGATCG 6480
QY 6481 TAGGGCTTAGAACCTCAGCACAACGTGCGAGCGTTCGCCATCAACGATACACCA 6540
Db 6481 TAGGGCTTAGAACCTCAGCACAACGTGCGAGCGGAACGTTTCCCATCAACGATACACCA 6540
QY 6541 CGGGACTTGGCACACCCTCCCGCGGCCCAACTATTCCAGGGCGCTATGCGGGTGGCTG 6600

Db 6541 CGGACCTTGCACACCTCCCGCGGCCCAACTATTCCAGGGCGCTATGCGGGTGGCTG 6600
QY 6601 CTGAGGAGTACGTGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGATGACCA 6660
Db 6601 CTGAGGAGTACGTGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGATGACCA 6660
QY 6661 CTGACAAAGTAAAGTGCCTATGCGAGGTTCCGGCCCCCGGAATTTCTTACGAGGTGGATG 6720
Db 6661 CTGACAAAGTAAAGTGCCTATGCGAGGTTCCGGCCCCCGGAATTTCTTACGAGGTGGATG 6720
QY 6721 GAGTGGGGTTGCA CAGGTACGCTCCGGGTGCAAACTCTTCTACGGGAGACGTCACT 6780
Db 6721 GAGTGGGGTTGCA CAGGTACGCTCCGGGTGCAAACTCTTCTACGGGAGACGTCACT 6780
QY 6781 TCCAGGTGCGGCTCAACCACTTGTGCGGTGCGAGTCCCATGCGAGCGCCGACCGG 6840
Db 6781 TCCAGGTGCGGCTCAACCACTTGTGCGGTGCGAGTCCCATGCGAGCGCCGACCGG 6840
QY 6841 ACGTAAACAGTCTTACTTCCATGCTCACCAGTCCCTCCCACTTACAGCAGAGCGGTA 6900
Db 6841 ACGTAAACAGTCTTACTTCCATGCTCACCAGTCCCTCCCACTTACAGCAGAGCGGTA 6900
QY 6901 AGCGTAGGCTGGCTAGAGGTTCTCCCGCTCTTTAGCCAGCTCATCAGTACGATGTT 6960
Db 6901 AGCGTAGGCTGGCTAGAGGTTCTCCCGCTCTTTAGCCAGCTCATCAGTACGATGTT 6960
QY 6961 CTGCGCTTCTTTGAAGCGACATGCACTACCCACCATGACTCCCGGACGCTGACCTCA 7020
Db 6961 CTGCGCTTCTTTGAAGCGACATGCACTACCCACCATGACTCCCGGACGCTGACCTCA 7020
QY 7021 TCGAGGGCAACCTCTTGTGGCGGAGGATGGCGGAAAAATCACTCGCGTGGAGTCAG 7080
Db 7021 TCGAGGGCAACCTCTTGTGGCGGAGGATGGCGGAAAAATCACTCGCGTGGAGTCAG 7080
QY 7081 AGAATAAGGTAGTAATTTCTGCACTCTTTTCGAAACCGCTTTCAGCGAGGGGATGAGAGG 7140
Db 7081 AGAATAAGGTAGTAATTTCTGCACTCTTTTCGAAACCGCTTTCAGCGAGGGGATGAGAGG 7140
QY 7141 AGATATCCGTGCGCGGAGATCTGCGAAATCCAGGAAGTTCCTCAGCGTTGCCCA 7200
Db 7141 AGATATCCGTGCGCGGAGATCTGCGAAATCCAGGAAGTTCCTCAGCGTTGCCCA 7200
QY 7201 TATGGGCAACGCGGACTACAACTCTCCACTGCTAGAGTCTCTGGAAGGACCGGACTAG 7260
Db 7201 TATGGGCAACGCGGACTACAACTCTCCACTGCTAGAGTCTCTGGAAGGACCGGACTAG 7260
QY 7261 TCCCTCCGGTGGTACACGGATGCCCATTTGCCACCTTACCAAGGCTCCTCAATACCACTC 7320
Db 7261 TCCCTCCGGTGGTACACGGATGCCCATTTGCCACCTTACCAAGGCTCCTCAATACCACTC 7320
QY 7321 CACGAGAAAGAGGACGGTTCTCTGACAGAAATCCAAATGTCTTCTGCTTGGCGGAGC 7380
Db 7321 CACGAGAAAGAGGACGGTTCTCTGACAGAAATCCAAATGTCTTCTGCTTGGCGGAGC 7380
QY 7381 TCGCCACTTAAGACCTTCCGGTAGCTCCGGATCGTGGCGGTTGATAGCGGACCGGACCG 7440
Db 7381 TCGCCACTTAAGACCTTCCGGTAGCTCCGGATCGTGGCGGTTGATAGCGGACCGGACCG 7440
QY 7441 CCCTTCTGACTGCGCTCCGACGAGCGGTGCAAAAGGATCCGACGTTGAGTCTGCTCT 7500
Db 7441 CCCTTCTGACTGCGCTCCGACGAGCGGTGCAAAAGGATCCGACGTTGAGTCTGCTCT 7500
QY 7501 CCATGCCCCCTTTGAAGGGAGCGGGGACCCCGATCTCAGCAGCGGCTCTTGGTCTA 7560
Db 7501 CCATGCCCCCTTTGAAGGGAGCGGGGACCCCGATCTCAGCAGCGGCTCTTGGTCTA 7560
QY 7561 CCGTGAAGTGAAGGCTAGTGAAGTGTGCTGCTGCTCAATGTCTTATACGTGGACAG 7620
Db 7561 CCGTGAAGTGAAGGCTAGTGAAGTGTGCTGCTGCTCAATGTCTTATACGTGGACAG 7620
QY 7621 GCGCCCTGATACGCCATCGCTGCGAGGAAAGTAAGCTGCCATCAACCCGTTGAGCA 7680
Db 7621 GCGCCCTGATACGCCATCGCTGCGAGGAAAGTAAGCTGCCATCAACCCGTTGAGCA 7680

QY	7681	ACTTTTGTGGTCAACCAACATGGTCTAGCCCAACATCTCCCGACGGCAAGCCTCC	7740
Db	7681	ACTTTTGTGGTCAACCAACATGGTCTAGCCCAACATCTCCCGACGGCAAGCCTCC	7740
QY	7741	GGCAGAAGAGGTCACCTTTTGACAGATTGCAAGTCTGTGATGATCAATACCAGGAGCTAC	7800
Db	7741	GGCAGAAGAGGTCACCTTTTGACAGATTGCAAGTCTGTGATGATCAATACCAGGAGCTAC	7800
QY	7801	TCAAGGAGATGAAGGGCAAGGCGTCCACAGTTAAGGCTAAGCTTCTATCTATAGAGAGG	7860
Db	7801	TCAAGGAGATGAAGGGCAAGGCGTCCACAGTTAAGGCTAAGCTTCTATCTATAGAGAGG	7860
QY	7861	CTGCAAGCTGACGCCGCCACATTCGGCCMAATTCGCTATGGGCTATGGGCAAGGACG	7920
Db	7861	CTGCAAGCTGACGCCGCCACATTCGGCCMAATTCGCTATGGGCTATGGGCAAGGACG	7920
QY	7921	TCCGGAACCTATCCAGCAGGCGGTTAAACACATCCGCTCCGCTGGAGGACTTCTGG	7980
Db	7921	TCCGGAACCTATCCAGCAGGCGGTTAAACACATCCGCTCCGCTGGAGGACTTCTGG	7980
QY	7981	AAGACTGAAACACCAATTTGACACACCACTCATGSCAAAAGTGAGGTTTCTGCGTCC	8040
Db	7981	AAGACTGAAACACCAATTTGACACACCACTCATGSCAAAAGTGAGGTTTCTGCGTCC	8040
QY	8041	AACCAGAGAGGAGGCGCAAGCAGCTCGCTTATTCGATTCACAGACTCGGAGTTC	8100
Db	8041	AACCAGAGAGGAGGCGCAAGCAGCTCGCTTATTCGATTCACAGACTCGGAGTTC	8100
QY	8101	GTGTATGCGAAGATGGCCCTTTACGAGTGTCTCCACCTTCTCAGGCGGTGATGG	8160
Db	8101	GTGTATGCGAAGATGGCCCTTTACGAGTGTCTCCACCTTCTCAGGCGGTGATGG	8160
QY	8161	GCTCCTCATACGATTTCATATCTCCCCCAAGCAGCGGTCGAGTTCCTGGTGAATACCT	8220
Db	8161	GCTCCTCATACGATTTCATATCTCCCCCAAGCAGCGGTCGAGTTCCTGGTGAATACCT	8220
QY	8221	GGAAATCAAAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGAATCAACGG	8280
Db	8221	GGAAATCAAAAGAAATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGAATCAACGG	8280
QY	8281	TCACTGAGATGACATTCGTGTTGAGAGTCAATTTACCAATGTTGACTTGGTCAACGG	8340
Db	8281	TCACTGAGATGACATTCGTGTTGAGAGTCAATTTACCAATGTTGACTTGGTCAACGG	8340
QY	8341	AGGCCACAGAGGACATGAGTTCGTACAGAGCGCTTTTACATCGGGGTCCTCGACTA	8400
Db	8341	AGGCCACAGAGGACATGAGTTCGTACAGAGCGCTTTTACATCGGGGTCCTCGACTA	8400
QY	8401	ACTCAAAAGGGCAGAACTGCGGTTTATCGCGGTGCGCGCAAGTGGCGTGTGACGACTA	8460
Db	8401	ACTCAAAAGGGCAGAACTGCGGTTTATCGCGGTGCGCGCAAGTGGCGTGTGACGACTA	8460
QY	8461	GCTCGGTATACCCCTCAATGTTTACCTGAGAGGCTGCGAGCTGTCGAGCTGCAAGC	8520
Db	8461	GCTCGGTATACCCCTCAATGTTTACCTGAGAGGCTGCGAGCTGTCGAGCTGCAAGC	8520
QY	8521	TCCAGGACTGACGATGCTCGTGAAACGAGACGACCTTGTGTTATCTGTGAAAGCGCGG	8580
Db	8521	TCCAGGACTGACGATGCTCGTGAAACGAGACGACCTTGTGTTATCTGTGAAAGCGCGG	8580
QY	8581	GAAACCCAGGAGATGCGCGGCCCTTACGAGCCTTTCAGGAGGCTATGACTAGGTATTCGG	8640
Db	8581	GAAACCCAGGAGATGCGCGGCCCTTACGAGCCTTTCAGGAGGCTATGACTAGGTATTCGG	8640
QY	8641	CCCCCCCCGGGGATCCGCCCCAACAGAAATACGACCTGAGCTGATCAATCATGTTTCCT	8700
Db	8641	CCCCCCCCGGGGATCCGCCCCAACAGAAATACGACCTGAGCTGATCAATCATGTTTCCT	8700
QY	8701	CCAATGTGTAGTCCGCGACGATGCTGCGAAAGGGTATCTACTCTACCCCGTGACC	8760
Db	8701	CCAATGTGTAGTCCGCGACGATGCTGCGAAAGGGTATCTACTCTACCCCGTGACC	8760

QY	8761	CCACACCCCTTGCACGGGCTGGTGGGACAGCTAGACACACTCTCAATCAACTCTT	8820
Db	8761	CCACACCCCTTGCACGGGCTGGTGGGACAGCTAGACACACTCTCAATCAACTCTT	8820
QY	8821	GGCTAGGCAATATCATGTATGCGCCACCTATGGCAAGGATGATTCGTGATGACTC	8880
Db	8821	GGCTAGGCAATATCATGTATGCGCCACCTATGGCAAGGATGATTCGTGATGACTC	8880
QY	8881	ACTTTTCTCCATCTCTTAGCTCAAGACAACTTGAAGGCACTTGAAGGCACTTGTGATCT	8940
Db	8881	ACTTTTCTCCATCTCTTAGCTCAAGACAACTTGAAGGCACTTGTGATCTTGTGATCT	8940
QY	8941	ACGGGCTTGTCTACTCCATTGAGCCACTTGAACCTACCTCAGATCAATTGAACGACTCATG	9000
Db	8941	ACGGGCTTGTCTACTCCATTGAGCCACTTGAACCTACCTCAGATCAATTGAACGACTCATG	9000
QY	9001	GTCTTAGCCATTTACATCCACAGTTACTTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060
Db	9001	GTCTTAGCCATTTACATCCACAGTTACTTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060
QY	9061	GCCTCAGGAACTTGGGGTACCACCTTGGGAACCTGGAGACATCGGGCCAGAGTGTCC	9120
Db	9061	GCCTCAGGAACTTGGGGTACCACCTTGGGAACCTGGAGACATCGGGCCAGAGTGTCC	9120
QY	9121	GGCTAAGCTACTGTCCAGGGGGGAGCGCCCACTTGTGGCAGATACCTCTTTAACT	9180
Db	9121	GGCTAAGCTACTGTCCAGGGGGGAGCGCCCACTTGTGGCAGATACCTCTTTAACT	9180
QY	9181	GGGCACTAAGGACCAAGCTTAAACTCACTCCAAATCCCGGCGGTCCAGCTGGACTGT	9240
Db	9181	GGGCACTAAGGACCAAGCTTAAACTCACTCCAAATCCCGGCGGTCCAGCTGGACTGT	9240
QY	9241	CTGGCTGTTGCTGCTGTTACAGCGGGGAGACATATATACAGGCTGTCTGCTGCC	9300
Db	9241	CTGGCTGTTGCTGCTGTTACAGCGGGGAGACATATATACAGGCTGTCTGCTGCC	9300
QY	9301	GACCCCGCTGTTTCCGTTGCTCTACTCTCTCTGTTAGGGTAGGCAATTTACCTGC	9360
Db	9301	GACCCCGCTGTTTCCGTTGCTCTACTCTCTCTGTTAGGGTAGGCAATTTACCTGC	9360
QY	9361	TCCCAACCCGATGAACGGGAGCTAACCACTCCAGGCTTTAAGCCATTTCTGTTTTT	9420
Db	9361	TCCCAACCCGATGAACGGGAGCTAACCACTCCAGGCTTTAAGCCATTTCTGTTTTT	9420
QY	9421	TT	9480
Db	9421	TT	9480
QY	9481	TTTTTCCCTTCTTTAATGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGTGTGAAAGGT	9540
Db	9481	TTTTTCCCTTCTTTAATGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGTGTGAAAGGT	9540
QY	9541	CCGTGAGCCGATGACTGACAGAGAGTGTGATCTAGTGGCTCTCTGACAGATCATGT	9595
Db	9541	CCGTGAGCCGATGACTGACAGAGAGTGTGATCTAGTGGCTCTCTGACAGATCATGT	9595

RESULT 4
AF054248
LOCUS Hepatitis C virus HC-J4, pCV-J4L2S, complete genome.
DEFINITION
ACCESSION AF054248
VERSION AF054248.1 GI:3098634
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.
REFERENCE 1 (bases 1 to 9595)
AUTHORS Yanagi,M., St Claire,M., Shapiro,M., Emerson,S.U., Purcell,R.H. and
Bukh,J.
Transcripts of a chimeric cDNA clone of hepatitis C virus genotype
1b are infectious in vivo

JOURNAL	Virology 244 (1), 161-172 (1998)	ORIGIN	
PUBMED	9581788	Query Match	99.2%; Score 9518.2; DB 13; Length 9595;
REFERENCE	2 (bases 1 to 9595)	Best Local Similarity	99.5%; Pred. No. 0;
AUTHORS	Yanagi,M. and Bukh,J.	Matches 9547; Conservative	0; Mismatches 48; Indels 0; Gaps 0;
TITLE	Direct Submision		
JOURNAL	Submitted (18-MAR-1998) Hepatitis Viruses Section, LID, NIAID, NIH,		
FEATURES	Dr., Building 7, Room 201, Bethesda, MD 20892, USA		
source	Location/Qualifiers		
	1..9595	Qy	1 GCCAGCCCCCTGATGGGGGACACTCCACCATGAATCACTCCCTCTGTGAGGAACACTACTG 60
	/organism="Hepatitis C virus"	Db	1 GCCAGCCCCCTGATGGGGGACACTCCACCATGAATCACTCCCTCTGTGAGGAACACTACTG 60
	/mol_type="genomic RNA"	Qy	61 TCCTTCACGAGAAAGCGTCTAGCCATCGCGTGTAGTATGATGATGCTGTGAGGCTCTAGGAC 120
	/strain="HC-J4"	Db	61 TCCTTCACGAGAAAGCGTCTAGCCATCGCGTGTAGTATGATGATGCTGTGAGGCTCTAGGAC 120
	/isolate="pcv-J4L2S"	Qy	121 CCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACGGTGAGTACACCGGAATTCGCAG 180
	/db_xref="taxon:11103"	Db	121 CCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACGGTGAGTACACCGGAATTCGCAG 180
	/note="non-infectious full-length cDNA"	Qy	181 GACGACCGGGTCCCTTTCTTGGATCAACCCGCTCAATGCTCGAGATTTGGCGCTGCCCCC 240
	342..9374	Db	181 GACGACCGGGTCCCTTTCTTGGATCAACCCGCTCAATGCTCGAGATTTGGCGCTGCCCCC 240
	/codon_start=1	Qy	241 GCGAGACTGTAGCGGAGTAGTGTGGTGGCGAAAGCCCTTGTGATCTGCTGATAGG 300
	/product="polyprotein"	Db	241 GCGAGACTGTAGCGGAGTAGTGTGGTGGCGAAAGCCCTTGTGATCTGCTGATAGG 300
	/protein_id="AAC15723.1"	Qy	301 GTGCTTGGAGTGGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTCTAAC 360
	/db_xref="GI:3098635"	Db	301 GTGCTTGGAGTGGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTCTAAC 360
	/translation="MSTNPQRKTRNTRRPQDKFPGGQIVGGVYLLPRGPRL	Qy	361 CTAAAGAAAAACCAACGTAACCAACCCGCGCCACACGAGCGTCAAGTTCCTCCGGCG 420
	GRVTRKASERQPRKQIPKARRPEGRAGWQGYEPWLYEGNEGUGAGWLLSPRG	Db	361 CTAAAGAAAAACCAACGTAACCAACCCGCGCCACACGAGCGTCAAGTTCCTCCGGCG 420
	SRPSPGTPRRSRNLKQVITLTCGFADLMGYIPLGAPLGGARALAHGVRLED	Qy	421 GTGTCAGATCGTTGGTGGAGTTTACCTGTGGCGCGCAGGGGCCCAAGTTCCTGATG 480
	GVNATGNLPGCSIFILLALISCLTIPASAYEVNRVSGIYHVTDLNCSNIVFEAAD	Db	421 GTGTCAGATCGTTGGTGGAGTTTACCTGTGGCGCGCAGGGGCCCAAGTTCCTGATG 480
	VTMTPGCPVCQVEGNSRCKWALPTTLAARNASVPTTIRRHVDLLVGTAAFCASMY	Qy	481 CGGCGACTAGGAAGGCTTCCGAGCGGTGCGAACCTCTGTGGAAGGGGACAACTATCCCA 540
	VGDLGCSIFLVSOLFTEPRRHEVTQDYNCSLYPCHVSHBMDWMNWSPTTALVV	Db	481 CGGCGACTAGGAAGGCTTCCGAGCGGTGCGAACCTCTGTGGAAGGGGACAACTATCCCA 540
	SOLLIPQAVDMVAGAHGVLGALAYYSVMGNWAKULIALLFAGVDETHTRGVA	Qy	541 AGGCTCGCGGACCCGAGGGAGGCGCTGGGCTCAGCCCGGGTACCTTGGCGCCCTCTATG 600
	GHTTSGFTSLNFGASQILVNTNGSHWIRNTALNCNDSLOTGFFAALFYAHKFNSS	Db	541 AGGCTCGCGGACCCGAGGGAGGCGCTGGGCTCAGCCCGGGTACCTTGGCGCCCTCTATG 600
	GCPEMASCRPTDWFQAGMPITYTKPNSDQRPYCMHYAPRCQVWPASQVCGPYG	Qy	601 GCAATGAGGGCTTGGCGGCTGAGGATGGCTCTGTGTACCCCGGGCTCCCGGCTAGTT 660
	FTPSPVAVGTTRDSGVPTYSWGENETDMLNNTRPQGNWFGCTWMNSTGFTKTCGG	Db	601 GCAATGAGGGCTTGGCGGCTGAGGATGGCTCTGTGTACCCCGGGCTCCCGGCTAGTT 660
	PCPNIGVGNRTLICPTDCFRKHPEATVTKGSGPMLTPRCLVDYPYRLWHYPCATLNF	Qy	661 GGGGCCCCACCGACCCCGGGTAGGTGCGTAACCTTGGGTAAAGTCAATCGATACCCCTTA 720
	SIFKVMYVGGVHELRNACNWRGECRNLDRDPSSELSPLLLSTEWQIILPCAFTTL	Db	661 GGGGCCCCACCGACCCCGGGTAGGTGCGTAACCTTGGGTAAAGTCAATCGATACCCCTTA 720
	PALSTGLHLHQNIYDVQILYGVGFVSFAIKWEYIILLFILLADARVCACLWNMLL	Qy	721 CATGGGCTTCGCGCATCTCATGGGTACATTCGCTCGTGGCGCCCCCTTAGGGGCG 780
	IAQAALLENLVNMAASVAGAGHILSLVFFCAAWYIKRGLAPGAAYFYFWPMLLL	Db	721 CATGGGCTTCGCGCATCTCATGGGTACATTCGCTCGTGGCGCCCCCTTAGGGGCG 780
	LLALLPFRAYALDRMAASCGGAVLGVFLTSLPYKFLVFLRLIWLQIYFTRAEAH	Qy	781 CTGCCAGGGCTTGGCACACCGGTGTCCGGGTCTTGGAGGACGGCGTGAACATATGCAACAG 840
	MOVVPLNVRGRDAIILLTCVAPPELIFDITKLLAILGLPMLVLAQIRVYFVR	Db	781 CTGCCAGGGCTTGGCACACCGGTGTCCGGGTCTTGGAGGACGGCGTGAACATATGCAACAG 840
	AOGLIRACMLVRKAGVGHYIOWAFMKIGALTCYVYNHLTPLRDWAHAGRLDVAVE	Qy	841 GGAATTCGCGGTTGCTCTTCTCTATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
	PVFSAMETKWTGADTAACGDIILGLPVSARGKEIFLGPADSLGQGRHLLAPIT	Db	841 GGAATTCGCGGTTGCTCTTCTCTATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
	ATSQGTGVLGVTITSLTRQKQVEGEVQVVSATQSFATCINGVCWTVIHGKSG	Qy	901 TCCAGCTTCGCGTTATGAAGTGGCAACGTGTCCGGGATATACCATGTCTACGAAAGCT 960
	TLAGPKPITQMTYNTLDLVGMQPPGARSNTPCCSGSDLLVLTTRADYIPVRRRG	Db	901 TCCAGCTTCGCGTTATGAAGTGGCAACGTGTCCGGGATATACCATGTCTACGAAAGCT 960
	DSRGLTNPSPVYSLKSGGGLPCPSGVHGVFRAAIVCTRGAAVDFIIPVESMETT	Qy	961 GCTCAACTCAAGCATTTGTGTATGAGGACGGGACGTGATCATGATATCTCCCGGGTGG 1020
	MRSPTVNTDNPSPVQVSOVAHLHAPTSGSKSTKPAAYAAQYKVLVLPVSAATL	Db	961 GCTCAACTCAAGCATTTGTGTATGAGGACGGGACGTGATCATGATATCTCCCGGGTGG 1020
	GFQYMSKAGHIDPMIRTVRTITTTGGSTYTSYCKFLADGCGSGGADIIICDECHS		
	TDSTIILGTLDOAETAGARLVLATATPPGSVTPHPNIEELGLSNGEIIPYK		
	ALPIEAKGGRHLIFCHSKKCDLAAKLTGLNVAAYIRGLDVSVPPIGDVVVVA		
	TDALMTGTFDSDIDNCVTQVDFSLDFTTETTTTPQDAVSRSQRRTGRG		
	RSGIYRFTVTPGSRPMDSSVLCYDAGCAWYELTPTAETSVRLRAYLNTPLGPVQC		
	DHLEWESVFTGLTHIDAFHSQTKQAGNFPYVAYOATVCARAQAPPSPWDQWKC		
	LIRLPTLHGPTLLYRLGAVONEVILTHPTIKYIMACMSADLEVVTSTWVPVGVLA		
	ALAAVCLTGSVVIIGRIILSKPAPVPDREVLYQEFDEMECASQLPYIISGMQLAE		
	QFKQKALGLLQATQKQAEAAAPVSRWRALETFWAKHMNFIISGQYLAGLSTLPGN		
	PAIASLMATFASITSPLTQHTLLNIGWVAQAQAPPAASAFAVGAAGVSGI		
	GLGKVLVDILAGVAGALVAFKVMGEVSTEDLVNLLPAILSPGALVGVVCAA		
	ILRRHVGPGEAVQWNRNLIAPASRGNHVSPTHYPESDAAARVTOILLSLTITOLK		
	RLHOMNEDCSTPCSGWLRYDWDICTVITDFKWLQSKLLPRLPGVPLFSCQGYK		
	GWRGDGMQTTCPGCAQIAGHVKNGSMRIUGPRTCSNTHWGTFFINATYTPGCPSP		
	APNYGRALVRAAEETVEVRVDHYVTGMTDNVCPCQVPAPEFTEVDVGRILHR		
	YAPACKLLREDVTQVGNLYLVGSLPCEPEPDVTVLTSMLTSTPHTAETAKRL		
	ARGSPSLASSASQISAPSLKATCTTHDSDPADLIEANLILWEMOENITRVSEN		
	KVVLIDSPPLHAEGBDEREISVAABELRKSRFPSPALPIWARPDYNNPILLESWKDPD		
	VPPVHVGCLPPTKAPPIPPRRKRTVLVTESNVSSALAEATKTFGSGSANDGT		
	ATALPDASDDKGDGSESYSMPLPEGPGDPLSDGSGVSTVSEBASSEDVCCSMS		
	YWTGALITPCAAESKLPINPLNSLRHNNVYATTSASLRKQKRVLFDRLQVLD		
	DHYRDLKEMKASTVAKLLSIEBAKLTTPHSAKSKFGYKADKVNRLSSRVNHI		
	RSVWEDLLETPTDITTIMAKSEVFCVQPEKGRKPARLIIVFPDLGVFRVCEKMALYD		
	VYSTLPOAVMGGSYGFQYSPKORVFLVNTWKSKKCPMGFSYDTRCFDSTVTESDIRV		
	ESIVQCCDLAPBAQQAISLTERLYIGPLTNSKGONCQYBRCFASGVLTSCGNTL		
	TCYLKATAACRAKALQDCITMLVNDGDLVVICESAGTOEDAAALRAATEAMTRYAPPG		
	DPQPEYDLELITSSRNSVVAHDASGRKYYLTDPITPLARAWEATRHPTVNSWL		
	GNIIMVATLHYSRGEINRVASCLRLKGLVPLRTWRHARSVRAKLLSQSGRAATCGRYL		
	GLSAPTLHSYSGEINRVASCLRLKGLVPLRTWRHARSVRAKLLSQSGRAATCGRYL		
	FNWAVTKLKLTPIPAASQLDLSGWFVAGYSGDIIYHSLSRARPRWFPLCLLLSVGV		
	GIYLPNPR"		

QY	1021	TGCCCCGTGTTTCAAGAGGTAACAGCTCCCGTTGCTGGGTAGCGTCACTCCACGCTCG	1080
DB	1021	TGCCCCGTGTTTCAAGAGGTAACAGCTCCCGTTGCTGGGTAGCGTCACTCCACGCTCG	1080
QY	1081	CGGCCAGGAATGCGACGGTCCCACTACGACAAATACAGCCACGTCGACTTGTCTGTTG	1140
DB	1081	CGGCCAGGAATGCGACGGTCCCACTACGACAAATACAGCCACGTCGACTTGTCTGTTG	1140
QY	1141	GGACGGCTGCTTTCTGCTCGCTATATAGTGGGGATCTCTCGGATCTATTTTCTCTCG	1200
DB	1141	GGACGGCTGCTTTCTGCTCGCTATATAGTGGGGATCTCTCGGATCTATTTTCTCTCG	1200
QY	1201	TCTCCAGCTGTTTCACTTCTCGCTCGCGGATGAGACAGTGCAGGACTGCAACTGCT	1260
DB	1201	TCTCCAGCTGTTTCACTTCTCGCTCGCGGATGAGACAGTGCAGGACTGCAACTGCT	1260
QY	1261	CAATCTATCCCGGCATGATACAGGTACCGCATGGCTTGGGATATGATGAATCTGGT	1320
DB	1261	CAATCTATCCCGGCATGATACAGGTACCGCATGGCTTGGGATATGATGAATCTGGT	1320
QY	1321	CACCTACACAGCCCTAGTGGTGTGCGAGTTGCTCCGGATCCCAAGCTGTCTGGACA	1380
DB	1321	CACCTACACAGCCCTAGTGGTGTGCGAGTTGCTCCGGATCCCAAGCTGTCTGGACA	1380
QY	1381	TGCTGCGGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGCTACTATTTCCATGGTAGGA	1440
DB	1381	TGCTGCGGGGGCCCACTGGGGAGTCTTGGCGGGCTTGGCTACTATTTCCATGGTAGGA	1440
QY	1441	ACTGGGCTAAGGTTCTGATTTGGCGTACTCTTTGGCGGGTGAAGGAGACCCACA	1500
DB	1441	ACTGGGCTAAGGTTCTGATTTGGCGTACTCTTTGGCGGGTGAAGGAGACCCACA	1500
QY	1501	CGACGGGAGGGTGGCGGCCACACACCTCGGGTTACGTCCTTTCTCTCATCTGGGG	1560
DB	1501	CGACGGGAGGGTGGCGGCCACACACCTCGGGTTACGTCCTTTCTCTCATCTGGGG	1560
QY	1561	CGTCTCAGAAAATCCAGCTTGTGAATACCAACGGCAGCTGGCAATCAACAGGACTGCC	1620
DB	1561	CGTCTCAGAAAATCCAGCTTGTGAATACCAACGGCAGCTGGCAATCAACAGGACTGCC	1620
QY	1621	TAAATTCGAATGACTCCTCCTCAAACTGGGTTCTTTGGCGGGCTGTTTACGCACAAAT	1680
DB	1621	TAAATTCGAATGACTCCTCCTCAAACTGGGTTCTTTGGCGGGCTGTTTACGCACAAAT	1680
QY	1681	TCAACTCGTCCGGTGGCCGAGCGCATGCGACCTGCCGCCCATTTGACTGGTTCCGCC	1740
DB	1681	TCAACTCGTCCGGTGGCCGAGCGCATGCGACCTGCCGCCCATTTGACTGGTTCCGCC	1740
QY	1741	AGGGGTGGGGCCCCATCACCTATATAAGCCTAAACAGCTCGGATCAGAGGCTTTATGCT	1800
DB	1741	AGGGGTGGGGCCCCATCACCTATATAAGCCTAAACAGCTCGGATCAGAGGCTTTATGCT	1800
QY	1801	GGCAATACGGCCTCGACCGTGGTGTGTAACCCGGTGCAGGTGTGTTGTCAGTGT	1860
DB	1801	GGCAATACGGCCTCGACCGTGGTGTGTAACCCGGTGCAGGTGTGTTGTCAGTGT	1860
QY	1861	ATTGTTTCAACCCCAAGCCCTGTTGTTGGGGACCCAGTCCGTTCCGGTCCCTACGT	1920
DB	1861	ATTGTTTCAACCCCAAGCCCTGTTGTTGGGGACCCAGTCCGTTCCGGTCCCTACGT	1920
QY	1921	ATAGCTGGGGGGAATGAGACAGCTGATGCTCCTCAACAAACACGCGTCCGCCACAAG	1980
DB	1921	ATAGCTGGGGGGAATGAGACAGCTGATGCTCCTCAACAAACACGCGTCCGCCACAAG	1980
QY	1981	GCAACTGGTTCGGCTGTACATGGAATAGTACTGGGTTCACTAAGACGTCGCGAGGTC	2040
DB	1981	GCAACTGGTTCGGCTGTACATGGAATAGTACTGGGTTCACTAAGACGTCGCGAGGTC	2040
QY	2041	CCCCGTGTAACTCGGGGGTTCGGTACCGCACTTGTATCTGCCCAACGACGCTTCC	2100
DB	2041	CCCCGTGTGTAACTCGGGGGTTCGGTAAACCGCACTTGTATCTGCCCAACGACGCTTCC	2100

QY	2101	GGAAACACCCCGAGGCTACTTACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT	2160
DB	2101	GGAAACACCCCGAGGCTACTTACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGGT	2160
QY	2161	GCCTAGTAGACTACCCATAACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCT	2220
DB	2161	GCCTAGTAGACTACCCATAACAGGCTTTGGCACTACCCCTGCACTCTCAATTTTTCATCT	2220
QY	2221	TTAAGGTTAGGATGTATGTGGGGGGTGGAGCACAGGCTCAATGCCGATGCAATTGGA	2280
DB	2221	TTAAGGTTAGGATGTATGTGGGGGGTGGAGCACAGGCTCAATGCCGATGCAATTGGA	2280
QY	2281	CTCGAGGAGCGCTGTAACTTTGGAGGACAGGATAGGTGAGAACTCAGCCCGCTGCTGC	2340
DB	2281	CTCGAGGAGCGCTGTAACTTTGGAGGACAGGATAGGTGAGAACTCAGCCCGCTGCTGC	2340
QY	2341	TGTCTACACAGAGTGGCAGATCTGCTGCTGCTTTTCAACCCCTACCGGCTTTATCCA	2400
DB	2341	TGTCTACACAGAGTGGCAGATCTGCTGCTGCTTTTCAACCCCTACCGGCTTTATCCA	2400
QY	2401	CTGGTTTGTATCCATCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGTGTAGGT	2460
DB	2401	CTGGTTTGTATCCATCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGTGTAGGT	2460
QY	2461	CAGCGTTTCTCTCTTTGCAATCAAAATGGAGTACATCTCTGTTGCTTTCTCTCTCTGG	2520
DB	2461	CAGCGTTTCTCTCTTTGCAATCAAAATGGAGTACATCTCTGTTGCTTTCTCTCTCTGG	2520
QY	2521	CAGACCGCGCGTGTGCTGCTGCTGCTGATGCTGCTGATAGCCAGGCTGAGGCCG	2580
DB	2521	CAGACCGCGCGTGTGCTGCTGCTGCTGATGCTGCTGATAGCCAGGCTGAGGCCG	2580
QY	2581	CCTTAGAGAACTTTGTTGTTCTCAATTCGCGCGTCCGTTGGCGGAGCGCATGTTCTCT	2640
DB	2581	CCTTAGAGAACTTTGTTGTTCTCAATTCGCGCGTCCGTTGGCGGAGCGCATGTTCTCT	2640
QY	2641	CTTTCTTGTGTTCTTCTGCGCGCTGTGATTAAGGCGAGCTGGCTCTCTGGGCGG	2700
DB	2641	CTTTCTTGTGTTCTTCTGCGCGCTGTGATTAAGGCGAGCTGGCTCTCTGGGCGG	2700
QY	2701	CGTATGCTTTTATGCGGTATGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2760
DB	2701	CGTATGCTTTTATGCGGTATGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2760
QY	2761	CTTACGCTTTGACACCGGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820
DB	2761	CTTACGCTTTGACACCGGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820
QY	2821	TATTTCTGACCTTGTCAACCATCTACAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCT	2880
DB	2821	TATTTCTGACCTTGTCAACCATCTACAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCT	2880
QY	2881	AATACCTTTATCACAGAGCGGAGCGCACATGCAAGTGTGGGTCTCCCGCTCAACGTTT	2940
DB	2881	AATACCTTTATCACAGAGCGGAGCGCACATGCAAGTGTGGGTCTCCCGCTCAACGTTT	2940
QY	2941	GGGAGGCGCGGATGCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3000
DB	2941	GGGAGGCGCGGATGCGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3000
QY	3001	ACATCACAAACT	3060
DB	3001	ACATCACAAACT	3060
QY	3061	CGAGAGTGGCTATCT	3120
DB	3061	CGAGAGTGGCTATCT	3120
QY	3121	AAGTTCGCGGGGGTCAATTATGTTCCAAATGGTCTTTCAAGAGTGGGCGCGCTGACGTA	3180
DB	3121	AAGTTCGCGGGGGTCAATTATGTTCCAAATGGTCTTTCAAGAGTGGGCGCGCTGACGTA	3180
QY	3181	CGTACGTTTATACCACTCTTACCCCACTGCGGGHACTGGGCCCAACGCGGGCTACGAGACC	3240

Db 3181 |||||CGTACGTTTATAAACCATCTTACCCCACTCGGGAGCTGGGCCCAACGCGGGCCTACGAGACC3240
QY 3241 TTGCGGTCGGGTAGAGCCGTCGTCTTCTCGGCCATCGAGACCAAGTGTATCATCCTGGG3300
Db 3241 TTGCGGTCGGGTAGAGCCGTCGTCTTCTCGGCCATCGAGACCAAGTGTATCATCCTGGG3300
QY 3301 GAGCAGACACCGCTCGCTGTGGGGACATCATCTTGGGTCTACCCGTCCTCGGCCGCAAGGG3360
Db 3301 GAGCAGACACCGCTCGCTGTGGGGACATCATCTTGGGTCTACCCGTCCTCGGCCGCAAGGG3360
QY 3361 GGAAGGAGATATTTTGGGACCGGTGATAGTCTCGAAGGCAAGGGTGGGACTCTCTTG3420
Db 3361 GGAAGGAGATATTTCTGGGACCGGTGATAGTCTCGAAGGCAAGGGTGGGACTCTCTTG3420
QY 3421 CGCCCATACCGCCTACTCCCAACAAAACGCGGGGCGTACTGGTGGCATCATCACTAGCC3480
Db 3421 CGCCCATACCGCCTACTCCCAACAAAACGCGGGGCGTACTGGTGGCATCATCACTAGCC3480
QY 3481 TCACAGGCGGACAAAGAACCAAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC3540
Db 3481 TCACAGGCGGACAAAGAACCAAGGTCGAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC3540
QY 3541 AATCTTTCTCGCGACCTGTCATCAACGCGGTGTGCTGGAATGTCTPACCATGCGCTGGCT3600
Db 3541 AATCTTTCTCGCGACCTGTCATCAATGCGGTGTGCTGGAATGTCTPACCATGCGCTGGCT3600
QY 3601 CGAAGACCTAGCCGTCGAAAGGTCCAAATCACCACCAATGTACACCAATGTAGACCTGG3660
Db 3601 CGAAGACCTAGCCGTCGAAAGGTCCAAATCACCACCAATGTAGACCTGG3660
QY 3661 ACCTGTCGGCTGGCAGGCGCCCGCGGGCGGCTCCATGACACCATGAGCTGTGGCA3720
Db 3661 ACCTGTCGGCTGGCAGGCGCCCGCGGGCGGCTCCATGACACCATGAGCTGTGGCA3720
QY 3721 GCTCGGACCTTTACTTTGGTCAGAGACATGCTGATGTCAATCCGGTGGCGCGGAGGCG3780
Db 3721 GCTCGGACCTTTACTTTGGTCAGAGACATGCTGATGTCAATCCGGTGGCGCGGAGGCG3780
QY 3781 ACAGAGGGGAAGTCTACTCTCCCGAGCCCGCTCTCTTACCTGAAAGCTCCTCGGGTG3840
Db 3781 ACAGAGGGGAAGTCTACTCTCCCGAGCCCGCTCTCTTACCTGAAAGCTCCTCGGGTG3840
QY 3841 GTCCATTGCTTTGCGGTCGCGGACGTCGTGGGGTCTTCCGGGCTGTGTGTCACACC3900
Db 3841 GTCCATTGCTTTGCGGTCGCGGACGTCGTGGGGTCTTCCGGGCTGTGTGTCACACC3900
QY 3901 GGGGGTCGCGAAGCGGTGGACTTCATACCCGTTGAGTCTATGGAAACTACCATGCCGT3960
Db 3901 GGGGGTCGCGAAGCGGTGGACTTCATACCCGTTGAGTCTATGGAAACTACCATGCCGT3960
QY 3961 CTCGGTCTTTCAGACAACTCAACCCCGCGGCTGTACCGAGACATTCGAAGTGGCNC4020
Db 3961 CCGCGTCTTTCAGACAACTCATCCCCCGGCTGTACCGAGACATCCCCAAGTGGCAC4020
QY 4021 ATCTGACGCTCTACTGTCGAGCGGCAAGACACAAAGTGGCGGTGGTATGAGCC4080
Db 4021 ATCTGACGCTCTACTGTCGAGCGGCAAGACACAAAGTGGCGGTGGTATGAGCC4080
QY 4081 AAGGGTACAAAGTGTCTGCTGAAACCCGTCGGTGGCGCCACCTTAGGGTTTGGGGCGT4140
Db 4081 AAGGGTACAAAGTGTCTGCTGAAACCCGTCGGTGGCGCCACCTTAGGGTTTGGGGCGT4140
QY 4141 ATATGTCCAAAGGCACACCGTATCGACCTTAAACATCAGAACTGGGGTAAGGACATTAACA4200
Db 4141 ATATGTCCAAAGGCACACCGTATCGACCTTAAACATCAGAACTGGGGTAAGGACATTAACA4200
QY 4201 CGGGCGGCTCCATTACGTAACCTATGGCAAGTTCCTTCGCAAGCGGTGGTCTTG4260
Db 4201 CGGGCGGCTCCATTACGTAACCTATGGCAAGTTCCTTCGCAAGCGGTGGTCTTG4260
QY 4261 GGGCGGCTTATGACATCATTAATATGTGATGAGTGGCCACTCAACTGACTCGACTACCATCT4320

Db 4261 GGGCGGCTTACGACATCATTAATATGTGATGAGTGCACCTCAACTGACTCGACTACCATCT4320
QY 4321 TTGSCATCGGCACAGCTCTCGACCAAGCGAGACGCTGGAGCGCGCTCGCTCGTGTCTCG4380
Db 4321 TTGSCATCGGCACAGCTCTCGACCAAGCGAGACGCTGGAGCGCGCTCGCTCGTGTCTCG4380
QY 4381 CCACCGCTCACCTTCGGGATCGGTTACCGGTGCGCACCCCCAATATCGAGGAAATAGGCC4440
Db 4381 CCACCGCTCACCTTCGGGATCGGTTACCGGTGCGCACCCCCAATATCGAGGAAATAGGCC4440
QY 4441 TGTCCAAATATGGAGATCCCTTCTATGGCAAGGCATCCCCATTTGAGGCGATCAAG4500
Db 4441 TGTCCAAATATGGAGATCCCTTCTATGGCAAGGCATCCCCATTTGAGGCGATCAAG4500
QY 4501 GGGGGAGGCATCTCATTTTCTGCCATTCGAAGAGAAATGTGACGAGCTCGCCGCAAGC4560
Db 4501 GGGGGAGGCATCTCATTTTCTGCCATTCGAAGAGAAATGTGACGAGCTCGCCGCAAGC4560
QY 4561 TGACAGGCTCGGACTGAACGCTGTAGCATATTAACCGGGGCTTTGATGTGTCCGTCTATC4620
Db 4561 TGACAGGCTCGGACTGAACGCTGTAGCATATTAACCGGGGCTTTGATGTGTCCGTCTATC4620
QY 4621 CGCCTATCGGAGACGTCGTTGTGCTGGCAACAGACGCTCTAATGACGGGTTTCAACGGCG4680
Db 4621 CGCCTATCGGAGACGTCGTTGTGCTGGCAACAGACGCTCTAATGACGGGTTTCAACGGCG4680
QY 4681 ATTTTGTACTGATGATCGACTGCAATACATGTGTCAACGACAGCTCGACTTCAGCTTGG4740
Db 4681 ATTTTGTACTGATGATCGACTGCAATACATGTGTCAACGACAGCTCGACTTCAGCTTGG4740
QY 4741 ATCCACCTTCCACATTTGAGACGACGCGTGCCTCCCAAGACGCGGTGTGCGCTCGCAAC4800
Db 4741 ATCCACCTTCCACATTTGAGACGACGCGTGCCTCCCAAGACGCGGTGTGCGCTCGCAAC4800
QY 4801 GGCAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCAGGAGAAC4860
Db 4801 GGCAGGTAGAACTGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCAGGAGAAC4860
QY 4861 GGCCCTCGGGCATGTCGATTTCTTGGTCTGTGAGTGTATGACGCGGGCTGTGCTT4920
Db 4861 GGCCCTCGGGCATGTCGATTTCTTGGTCTGTGAGTGTATGACGCGGGCTGTGCTT4920
QY 4921 GGTATGAGCTCACGCCGCTGAGACCTCGGTTAGGTTTGGGGCTTACCTAAATACACCG4980
Db 4921 GGTATGAGCTCACGCCGCTGAGACCTCGGTTAGGTTTGGGGCTTACCTAAATACACCG4980
QY 4981 GGTTCGCGCTGCGCAGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGCTCAACC5040
Db 4981 GGTTCGCGCTGCGCAGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGCTCAACC5040
QY 5041 ACATAGATGCCACTTCCCTGTCAGACTAAACAGGCAAGGAGACAACTTTCCTTACCTGG5100
Db 5041 ACATAGATGCCACTTCCCTGTCAGACTAAACAGGCAAGGAGACAACTTTCCTTACCTGG5100
QY 5101 TGGCATATCAAGCTACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCTACGTTGGGACCAA5160
Db 5101 TGGCATATCAAGCTACAGTGTGCGCCAGGGCTCAAGCTCCACCTCCTACGTTGGGACCAA5160
QY 5161 TGTGAAAGTGTCTCATCGGCTGAAACCTACACTGACGCGGCAACACCCCTGCTGTATA5220
Db 5161 TGTGAAAGTGTCTCATCGGCTGAAACCTACACTGACGCGGCAACACCCCTGCTGTATA5220
QY 5221 GGCTAGGAGCGCTCCAAATGAGGTCTCTCACACCCCATTAATTAATCATCATGG5280
Db 5221 GGCTAGGAGCGCTCCAAATGAGGTCTCTCTCACACCCCATTAATTAATCATCATGG5280
QY 5281 CATGCAATGTCCGCTGACCTCGGAGTGTGCTAGCAGCTGGGTGTGCTAGCGGAGTCC5340
Db 5281 CATGCAATGTCCGCTGACCTCGGAGTGTGCTAGCAGCTGGGTGTGCTAGCGGAGTCC5340
QY 5341 TTGCAAGCTTTGGCGGCATCTGCTGACGACAGGCGAGTGTGCTCATTTGGGCGAGATCA5400
Db 5341 TTGCAAGCTTTGGCGGCATCTGCTGACGACAGGCGAGTGTGCTCATTTGGGCGAGATCA5400

QY	5401	TCCTGTCCGGGAAGCCAGCTGTGCTTCCCGACAGGGAAGTCTCTACACAGGAGTTCGATG	5460
Db	5401	TTTGTCCGGGAAGCCAGCTGTGCTTCCCGACAGGGAAGTCTCTACACAGGAGTTCGATG	5460
QY	5461	AGATGGAAGAGTGTGCTCAACACTTCTTACATCGACGAGGGAATGCGCTCGCGAGC	5520
Db	5461	AGATGGAAGAGTGTGCTCAACACTTCTTACATCGACGAGGGAATGCGCTCGCGAGC	5520
QY	5521	AAATTCAAGCAAAAGGGCGCTCGGCTTGTGCAACCGGCCAACAGCAAGCGAGCGTCTG	5580
Db	5521	AAATTCAAGCAAAAGGGCGCTCGGCTTGTGCAACCGGCCAACAGCAAGCGAGCGTCTG	5580
QY	5581	CTCCGCTGTGTGAGTCAAGTGTGCGAGCCCTTGAGACCTTCTGTGGGGAAGCAATGTGGA	5640
Db	5581	CTCCGCTGTGTGAGTCAAGTGTGCGAGCCCTTGAGACCTTCTGTGGGGAAGCAATGTGGA	5640
QY	5641	ATTTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCTGGAACCCCGCGA	5700
Db	5641	ATTTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCTGGAACCCCGCGA	5700
QY	5701	TAGCATCATTTGATGGCAATTTACAGCTTCTATCACTAGCCCGCTCACACCCCAAAACACCC	5760
Db	5701	TAGCATCATTTGATGGCAATTTACAGCTTCTATCACTAGCCCGCTCACACCCCAAAACACCC	5760
QY	5761	TCCTGTTTAAACATCTTGTGGGGGATGTGGTGTGCTGCCAACTCGCTCTCTCCAGCGCTGCGT	5820
Db	5761	TCCTGTTTAAACATCTTGTGGGGGATGTGGTGTGCTGCCAACTCGCTCTCTCCAGCGCTGCGT	5820
QY	5821	CAGCTTTCGTGTGGCGCGGCATCGCCGAGGGCTGTGGCAGCATAGGCTTGTGGGAAG	5880
Db	5821	CAGCTTTCGTGTGGCGCGGCATCGCCGAGGGCTGTGGCAGCATAGGCTTGTGGGAAG	5880
QY	5881	TGCTCGTGGACATCTTGTGGGGCTATGGGGCAGGGGTAGCGGGCACTCGTGCCCTTTA	5940
Db	5881	TGCTCGTGGACATCTTGTGGGGCTATGGGGCAGGGGTAGCGGGCACTCGTGCCCTTTA	5940
QY	5941	AGGTCATGAGCGGCGAGGTGCCTTCCACCGAGGACCTTGGTCAACTTACTCCCTGCCATCC	6000
Db	5941	AGGTCATGAGCGGCGAGGTGCCTTCCACCGAGGACCTTGGTCAACTTACTCCCTGCCATCC	6000
QY	6001	TCTCTCTGTGTGCTGTGCTGGGCTGTGCTGGGCGAGCAATACTCGCTCGGCACTGG	6060
Db	6001	TCTCTCTGTGTGCTGTGCTGGGCTGTGCTGGGCGAGCAATACTCGCTCGGCACTGG	6060
QY	6061	GCCCGGAGGGGGCTGTGCAAGTGAACCGCTGTAGACGCTCGCTTCGGCGGTA	6120
Db	6061	GCCCGGAGGGGGCTGTGCAAGTGAACCGCTGTAGACGCTCGCTTCGGCGGTA	6120
QY	6121	ACCAAGCTCTCCCTACGCACTATGTGCTGAGAGCGAGCTGACAGCACTGCTCACTCAGA	6180
Db	6121	ACCAAGCTCTCCCTACGCACTATGTGCTGAGAGCGAGCTGACAGCACTGCTCACTCAGA	6180
QY	6181	TCCTCTCTAGCCTTACCATCACTCAACTGTGGAAGCGGCTCCACAGTGGATTAATGAGG	6240
Db	6181	TCCTCTCTAGCCTTACCATCACTCAACTGTGGAAGCGGCTTCCACAGTGGATTAATGAGG	6240
QY	6241	ACTGCTCTACGCACTGCTCGGCTGTGCTTAAGGATGTTTGGATTTGGATATGCAAGG	6300
Db	6241	ACTGCTCTACGCACTGCTCGGCTGTGCTTAAGGATGTTTGGATTTGGATATGCAAGG	6300
QY	6301	TGTTGACTGACTTCAAGACCTTGGCTCCAGTCCAAACTCCTCGCGGCTTACCGGGAGTCC	6360
Db	6301	TGTTGACTGACTTCAAGACCTTGGCTCCAGTCCAAACTCCTCGCGGCTTACCGGGAGTCC	6360
QY	6361	CTTTCTCTGTCATGCCAACCGGGTCAAGAGGAGTCTGCGGGGGAAGCGCATCATGCAAA	6420
Db	6361	CTTTCTCTGTCATGCCAACCGGGTCAAGAGGAGTCTGCGGGGGAAGCGCATCATGCAAA	6420
QY	6421	CCACTGCCCATGCGGAGCACAGATCGCGGAGCATGTCAAAAACGGTTCATGAGGATCG	6480
Db	6421	CCACTGCCCATGCGGAGCACAGATCGCGGAGCATGTCAAAAACGGTTCATGAGGATCG	6480

QY	6481	TAGGCGCTAGAACCTGCGACAAACAGTGGCAACGAAAGTTCCTCCCATCAACAGCATACACCA	6540
Db	6481	TAGGCGCTAGAACCTGCGACAAACAGTGGCAACGAAAGTTCCTCCCATCAACAGCATACACCA	6540
QY	6541	CGGGACCTTTGCACACCTCCCGGGGCCCAACTATTTCAGAGGCGCTATGCGGGTGGCTG	6600
Db	6541	CGGGACCTTTGCACACCTCCCGGGGCCCAACTATTTCAGAGGCGCTATGCGGGTGGCTG	6600
QY	6601	CTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGATGACCA	6660
Db	6601	CTGAGGAGTACGTGGAGGTTACGCGTGTGGGGATTTCCACTACGTGACGGGATGACCA	6660
QY	6661	CTGACAACTGAAGTGCCTCATGCTCCGCGCTCCGCGCTCCGAAATTTCTTCAAGAGGTGGATG	6720
Db	6661	CTGACAACTGAAGTGCCTCATGCTCCGCGCTCCGCGCTCCGAAATTTCTTCAAGAGGTGGATG	6720
QY	6721	GAGTGTGGTTGACACAGGTACGCTCCGCGCTGCACAACTCTTCTTACGGGAGGACGTCACTG	6780
Db	6721	GAGTGTGGTTGACACAGGTACGCTCCGCGCTGCACAACTCTTCTTACGGGAGGACGTCACTG	6780
QY	6781	TCAGGTCCGGCTCAACCAATACTTGGTTCGGTTCGAGCTCCCATGCGGAGCCGACCGG	6840
Db	6781	TCAGGTCCGGCTCAACCAATACTTGGTTCGGTTCGAGCTCCCATGCGGAGCCGACCGG	6840
QY	6841	ACGTAACAGTGTACTTCCATGCTCACCGATCCCTCCACATTCACAGACAGACCGCTA	6900
Db	6841	ACGTAACAGTGTACTTCCATGCTCACCGATCCCTCCACATTCACAGACAGACCGCTA	6900
QY	6901	AGCTAGGCTGGCTAGAGGCTCTCCCGCTCTTTAGCCAGCTCAATCAGCTAGGACGTTGT	6960
Db	6901	AGCTAGGCTGGCTAGAGGCTCTCCCGCTCTTTAGCCAGCTCAATCAGCTAGGACGTTGT	6960
QY	6961	CTCGGCTCTTTTGAAGGCGACATGACTACCAACCATGACTCCCGGAGCGCTGACCTCA	7020
Db	6961	CTCGGCTCTTTTGAAGGCGACATGACTACCAACCATGACTCCCGGAGCGCTGACCTCA	7020
QY	7021	TCAGGCGCAACCTCTTGTGGCGCAGAGATGGCGGAAACATCATCTCGGTCGAGTCAG	7080
Db	7021	TCAGGCGCAACCTCTTGTGGCGCAGAGATGGCGGAAACATCATCTCGGTCGAGTCAG	7080
QY	7081	AGATAAGGTAGTAATTTCTGGAATCTTTCGAAACCGCTTTCAGCGGAGGGGATGAGAGG	7140
Db	7081	AGATAAGGTAGTAATTTCTGGAATCTTTCGAAACCGCTTTCAGCGGAGGGGATGAGAGG	7140
QY	7141	AGATAATCCGTGCGGGGAGATCCTGCGAAATTCAGGAAAGTTCCTCCTCAGCGTGGCCA	7200
Db	7141	AGATAATCCGTGCGGGGAGATCCTGCGAAATTCAGGAAAGTTCCTCCTCAGCGTGGCCA	7200
QY	7201	TATGGGCAACCGCGGACTCAATCTCTCCTGCTAGAGTCTTGGAGGACCCGAGCTACG	7260
Db	7201	TATGGGCAACCGCGGACTCAATCTCTCCTGCTAGAGTCTTGGAGGACCCGAGCTACG	7260
QY	7261	TCCCTCGGTGTACACGGATGCCATTGCGACCTACCAAGGCTCTCTCCAATACCACTC	7320
Db	7261	TCCCTCGGTGTGTACACGGATGCCATTGCGACCTACCAAGGCTCTCTCCAATACCACTC	7320
QY	7321	CACGAGAAAGAGGACGGTTGTCTGTGACGAATCCAATGTCTTCTGCTTGGCGGAGC	7380
Db	7321	CACGAGAAAGAGGACGGTTGTCTGTGACGAATCCAATGTCTTCTGCTTGGCGGAGC	7380
QY	7381	TCGCGCACTAGACCTTCGGTAGCTCCGGATCGTCCGCGGTTGATAGCGGACCGGACCG	7440
Db	7381	TCGCGCACTAGACCTTCGGTAGCTCCGGATCGTCCGCGGTTGATAGCGGACCGGACCG	7440
QY	7441	CCCTTCTGACCTCGGCTCCGACGAGCGTGACAAAGGATCCGAGCTTGTGAGTCGTACTCCT	7500
Db	7441	CCCTTCTGACCTCGGCTCCGACGAGCGTGACAAAGGATCCGAGCTTGTGAGTCGTACTCCT	7500
QY	7501	CCATGCCCGCTTGAAGGGGAGCCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTA	7560
Db	7501	CCATGCCCGCTTGAAGGGGAGCCGGGGACCCCGATCTCAGCGACGGGTCTTGGTCTA	7560
QY	7561	CCGTGAGTGAAGGCTAGTGAGGATGCTGCTGCTCTCAATGCTCTATACGTGGACAG	7620

Db ||||| 7561 CCGTGAGTGAGGAGCTAGTGAGGACGTCGTCTGCTGCTCAATGTCTCTATACGTGGACAG 7620
QY ||||| 7621 GGGCCCTGATCAGCCATCGCTGCGGAGGAAGTAAGCTGCCCATCAACCGTTGACCA 7680
Db ||||| 7621 GGGCCCTGATCAGCCATCGCTGCGGAGGAAGTAAGCTGCCCATCAACCGTTGACCA 7680
QY ||||| 7681 ACTCTTTGCTGGTCCACACAAATGGTCTACGCCACAAATCCCGCAGCGCAAGCCCTCC 7740
Db ||||| 7681 ACTCTTTGCTGGTCCACACAAATGGTCTACGCCACAAATCCCGCAGCGCAAGCCCTCC 7740
QY ||||| 7741 GGCAGAAGAGTCACTTTGACAGATTGCAAGTCTCGATGATCAATACCGGGACGTAC 7800
Db ||||| 7741 GGCAGAAGAGTCACTTTGACAGATTGCAAGTCTCGATGATCAATACCGGGACGTAC 7800
QY ||||| 7801 TCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGGAG 7860
Db ||||| 7801 TCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGGAG 7860
QY ||||| 7861 CCGCAAGCTGACGCCCCACACATTCGGCCAAATCCAAATTTGGCTATGGGCAAGGACG 7920
Db ||||| 7861 CCGCAAGCTGACGCCCCACACATTCGGCCAAATCCAAATTTGGCTATGGGCAAGGACG 7920
QY ||||| 7921 TCCGAAACCTATCCAGCAGGCGGTTAAACCAATCCGCTCCGTTGGGAGGACTTTGCTGG 7980
Db ||||| 7921 TCCGAAACCTATCCAGCAGGCGGTTAAACCAATCCGCTCCGTTGGGAGGACTTTGCTGG 7980
QY ||||| 7981 AAGACACTGAACACCAATTTGACACCATCATGGCAAAAGTAGGAGTTTCTCGTCC 8040
Db ||||| 7981 AAGACACTGAACACCAATTTGACACCATCATGGCAAAAGTAGGAGTTTCTCGTCC 8040
QY ||||| 8041 AACCCAGAGGAGGCGCCAGACGCTCGCTTTATGTTATCCAGACCTGGGAGTTTC 8100
Db ||||| 8041 AACCCAGAGGAGGCGCCAGACGCTCGCTTTATGTTATCCAGACCTGGGAGTTTC 8100
QY ||||| 8101 GTGTATGCGAGAGATGGCCCTTTACGACGTGGTCTCCACCCCTTCTCAGGCCGTGATGG 8160
Db ||||| 8101 GTGTATGCGAGAGATGGCCCTTTACGACGTGGTCTCCACCCCTTCTCAGGCCGTGATGG 8160
QY ||||| 8161 GCTCCTCATACGATTTCAATCTCCCAAGCAGGCGGTGAGTTCCTGTGTAATACCT 8220
Db ||||| 8161 GCTCCTCATACGATTTCAATCTCCCAAGCAGGCGGTGAGTTCCTGTGTAATACCT 8220
QY ||||| 8221 GGAATCAAGAAATGCCCCCTATGAGGAGTCAATTTACCAATTTGTAAGTCAAGCG 8280
Db ||||| 8221 GGAATCAAGAAATGCCCCCTATGAGGAGTCAATTTACCAATTTGTAAGTCAAGCG 8280
QY ||||| 8281 TCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACCAATTTGTAAGTCAAGCG 8340
Db ||||| 8281 TCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACCAATTTGTAAGTCAAGCG 8340
QY ||||| 8341 AGGCCAGACAGGCATTAAGTTCGCTCAGAGCGGCTTTACATCGGGGTCGCCCTGACTA 8400
Db ||||| 8341 AGGCCAGACAGGCATTAAGTTCGCTCAGAGCGGCTTTACATCGGGGTCGCCCTGACTA 8400
QY ||||| 8401 ACTCAAAAGGCGAGACTCGCGTTATCGCGGTTACGCGCAAGTGGCGTGTGACGACTA 8460
Db ||||| 8401 ACTCAAAAGGCGAGACTCGCGTTATCGCGGTTACGCGCAAGTGGCGTGTGACGACTA 8460
QY ||||| 8461 GCTGCGGTAAATACCTTCACATGTTACTTGAAGGCCACCTGACGCTGTGAGCTGCAAAAGC 8520
Db ||||| 8461 GCTGCGGTAAATACCTTCACATGTTACTTGAAGGCCACCTGACGCTGTGAGCTGCAAAAGC 8520
QY ||||| 8521 TCAGGACTGACAGATGCTCGTGAACCGGAGACGACCTTCTGCTTATCTGTGAAAGCGCG 8580
Db ||||| 8521 TCAGGACTGACAGATGCTCGTGAACCGGAGACGACCTTCTGCTTATCTGTGAAAGCGCG 8580
QY ||||| 8581 GAACCCAGAGGATGCGGCGCCCTACGAGCTTCCAGGAGCTATGACTAGGTATTCG 8640
Db ||||| 8581 GAACCCAGAGGATGCGGCGCCCTACGAGCTTCCAGGAGCTATGACTAGGTATTCG 8640
QY ||||| 8641 CCCCCCGGGGATCCGCCCAACCAAGATAAGACCTGGAGCTGATAACATCATGTCTCT 8700

Db ||||| 8641 CCCCCCGGGGATCCGCCCAACCAAGATAAGACCTGGAGCTGATAAATACATCACGTCTCT 8700
QY ||||| 8701 CCAATGTGTCAGTCGCGCACGATGATCTGCAAAAGGGTATATCTACTCAACCCGTGACC 8760
Db ||||| 8701 CCAATGTGTCAGTCGCGCACGATGATCTGCAAAAGGGTATATCTACTCAACCCGTGACC 8760
QY ||||| 8761 CCACCAACCCCTTTGACAGGCGTCTGCTGGGAGACAGCTAGACACATCCCAATCAACTCTT 8820
Db ||||| 8761 CCACCAACCCCTTTGACAGGCGTCTGCTGGGAGACAGCTAGACACATCCCAATCAACTCTT 8820
QY ||||| 8821 GGTAGGCAATATCATGATATGCGGCCACCTATGGGCAAGGATGATTTGTATGATCTC 8880
Db ||||| 8821 GGTAGGCAATATCATGATATGCGGCCACCTATGGGCAAGGATGATTTGTATGATCTC 8880
QY ||||| 8881 ACTTTTCTCCATCTCTAGCTCAAGACAACTTTGAAAAGCCCTGATTTGTGATGATCT 8940
Db ||||| 8881 ACTTTTCTCCATCTCTAGCTCAAGACAACTTTGAAAAGCCCTGATTTGTGATGATCT 8940
QY ||||| 8941 ACGGGGCTTGCTACTCCATTTGAGCCACTTGACCTTACCTCAGATCAATTTGAACGACTCCATG 9000
Db ||||| 8941 ACGGGGCTTGCTACTCCATTTGAGCCACTTGACCTTACCTCAGATCAATTTGAACGACTCCATG 9000
QY ||||| 9001 GTCTTAGGCAATTTACATCTCCACAGTTACTCTCCAGGTGAGATCAATAGGTTGCTTCAT 9060
Db ||||| 9001 GTCTTAGGCAATTTACATCTCCACAGTTACTCTCCAGGTGAGATCAATAGGTTGCTTCAT 9060
QY ||||| 9061 GCCTCAGAAACTTTGGGTTACCCCTTTGGAACTCTGAGACATCGGGCCAGAAAGTGTC 9120
Db ||||| 9061 GCCTCAGAAACTTTGGGTTACCCCTTTGGAACTCTGAGACATCGGGCCAGAAAGTGTC 9120
QY ||||| 9121 GCGCTAAGCTACTCTCCAGGGGGAGGCGCCACTTTGTGGCAGATACCTCTTTAACT 9180
Db ||||| 9121 GCGCTAAGCTACTCTCCAGGGGGAGGCGCCACTTTGTGGCAGATACCTCTTTAACT 9180
QY ||||| 9181 GGGCAGTAAGAACCAAGCTTAACTCACTCAATCCCGCCGCGCTCCAGCTGCACTTGT 9240
Db ||||| 9181 GGGCAGTAAGAACCAAGCTTAACTCACTCAATCCCGCCGCGCTCCAGCTGCACTTGT 9240
QY ||||| 9241 CTGGCTGGTTCGTCGCTGTTTACAGCGGGGAGACATATATACAGCCTGCTCGTGCCC 9300
Db ||||| 9241 CTGGCTGGTTCGTCGCTGTTTACAGCGGGGAGACATATATACAGCCTGCTCGTGCCC 9300
QY ||||| 9301 GACCCCGCTGCTGCTGCTGCTTCTCTCTCTCTCTAGGGGTAGGCATTTACCTGTC 9360
Db ||||| 9301 GACCCCGCTGCTGCTGCTGCTTCTCTCTCTCTCTCTAGGGGTAGGCATTTACCTGTC 9360
QY ||||| 9361 TCCCAACCGATGAACCGGGAGCTTAAACCTCCAGGCGTTAAGCCATTTCTGTTTTTTT 9420
Db ||||| 9361 TCCCAACCGATGAACCGGGAGCTTAAACCTCCAGGCGTTAAGCCATTTCTGTTTTTTT 9420
QY ||||| 9421 TT 9480
Db ||||| 9421 TT 9480
QY ||||| 9481 TTTTTTCCCTTCTTTAAATGGTGGCTCCATCTTAGCCCTAGTCAACGGCTAGCTGTGAAAGT 9540
Db ||||| 9481 TTTTTTCCCTTCTTTAAATGGTGGCTCCATCTTAGCCCTAGTCAACGGCTAGCTGTGAAAGT 9540
QY ||||| 9541 CCGTAGCGCGCATGACTGACAGAGTGTGTATATGCGCTCTCTGACAGATCATGT 9595
Db ||||| 9541 CCGTAGCGCGCATGACTGACAGAGTGTGTATATGCGCTCTCTGACAGATCATGT 9595

RESULT 5

AF054249
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AF054249
Hepatitis C virus
AF054249
Hepatitis C virus
AF054249.1
GI:3098636
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

5956 bp
RNA
linear
VRL 02-MAY-1998
complete genome.

Hepacivirus.
1 (bases 1 to 9596)
Yanagi,M., St Claire,M., Shapiro,M., Emerson,S.U., Purcell,R.H. and
Bukh,J.

Transcripts of a chimeric cDNA clone of hepatitis C virus genotype
1b are infectious in vivo
Virology 244 (1), 161-172 (1998)

9581788

2 (bases 1 to 9596)

Yanagi,M. and Bukh,J.

Direct Submission

Submitted (18-MAR-1998) Hepatitis Viruses Section, LID, NIAID, NIH,

7 Center Dr., Building 7, Room 201, Bethesda, MD 20892, USA

Location/Qualifiers

1. .9596

/organism="Hepatitis C virus"

/mol_type="genomic RNA"

/strains="HC-J4"

/isolate="pCV-J4L4S"

/db_xref="taxon:11103"

/notes="non-infectious full-length cDNA"

343. .9375

/codon_start=1

/product="polyprotein"

/protein_id="AAC15724.1"

/db_xref="GI:3098637"

/translation="MSTNPKPQRKTKRNTNRPPQDKVFPFGGQIVGGVLLPRRGRPL
GVRAKRTSRSPRRORRPIPKARRPEGRAPQGPMPPLYGNEGLWAGMLLSPRG
SRPSPGTPTRRSRNLGVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVED
GVNYATGNLPGCSFSIPLKALLSLCTIPASAVEVRNVSGIYHVYNDNSNSIYVEAAD
VINHTPGCPVCREGNSRCWALPTTLAARDASVFTTIRRHVDLLVCSNTAFCSAMV
VGLDGSIFLVSOLFSPRRHTVDQNCISYPIGHVSGHRMAWMMNWSPTALVV
SOLLRTPOAVMVGAWGLVLAGLAYISVMGNWAKVLIVALLFAGVDGETHTGRV
GRTTGLTGLVLDQAEATAGARLVLATATPPGSVTPHPNIEEIGLSNNGEIPFYVK
TDSITTLIGITVLDOAETAGARLVLATATPPGSVTPHPNIEEIGLSNNGEIPFYVK
PALSTGLIHLQNVQVLYGVGSFVFAIKWEYILLLFLSLADARVACMLMMLL
IAQAEALLENVLVNAASVAGAGHLSFLVFFCAAWIKRGLAPGAPAYFVGPWPLL
LIALPRAVLDREMAASCGGAVLGLVPLTISPVYKVLPLRLIWLQYFIIRAEAH
LQGLIRACMLRVAGHVQMAKLGALTGYVYNHLTPLRDWAHAGURDLAAVE
MOVVPLNVRGDRDAIILITCAVHPELIFDITKLLALLGLPLMLVLAQGITRVYFVR
PVVFSAMETKVIITWADTAAACDIIILGLPSARRGEIIFLGPADSLGGQWRLLAPIT
AYSQOQRTGVLTSLTRDKNQVEGVQVSTATQSFATCINGVTVTVHGAAGSK
TLAAGPKGITOMTVNDLVLVQWAPGARSMTPCSCGSDLVLVTRHADVIPIVRERG
DSRGSLLSPPRVSLLKSGSGPLLCPSGHVGVGFRAAVCTRGVAKAVDFIPIVBSMETT
MRSPVFTDSSPAVQTFQVAHLHAPTGSKSTKUPAAVAAQYKVLNLPVSAATL
GFGAYMSKAGHDINRTGVRTITTGSGITSYTKFLADGGSGGAYDIIICDECHS
TDSITTLIGITVLDOAETAGARLVLATATPPGSVTPHPNIEEIGLSNNGEIPFYVK
AIPIEAIKGGRHILFCHSKKKCDLAAKLTGLGLNAVAYYRGLDVSIVPIPGDVVYA
TDALMTGCTGDSVDICNTCVTQVDFSLDPTFTITTTVPQDASVRSORRGTGRG
RSGIYRFTPGERSPGMSFVLCCEYDAGCAMHELTTPAETSRLRAYLNTPLGLPVCO
DHLFEVSFETGLTHIDAHFLSQTKQAGDNFFILVAYQATVCAQAAPPSPSQWMKWC
LIRLKPETHGTPFLLRVLAGVQNEVILTHPIKYIMACMSADLEVTSTWLVGGVLA
ALAAKCLTTSVTVIGRRIILSGRPVAVPDEVLYQEFDEMECASOLPYTEQGMQALB
PQKQALGLQATKQAEAAAPVVEKSRALFTFWAKMWNFIISGQYLAGSLTLPGN
PFAJLMAFTSITSPLTNTQNTLNLGLGWAVALAPPSAASAFAGAGTAGAAGVSI
GLKVLVDIILAGVAGVAGALVAFKVMGSEVPSTEDLVNLLPALLSPGALVGVVCA
LRRHVFGPBGAVQMMNLIAFASGRNHVSPTHVTPESDAAARVQLPSLTITQLIK
RLHQWINDCSTPCSGSMLRDVMDWICTVLTDFTKLQSLKLPRLPGLVPSLCQRYK
GVNRGDGIMOTPCGAGIAGHVKNRIVGPRCTNSNTWHGTFPIINAYTTPCTPSF
APNYSRALWRAAEYVEVTRVGDVHVVTGMTDNVCKPCQVPAPPEFTFVGDVRLHR
YAPACRPLLEDVTFQVGLNQYLVGSQLPCEPEPDDVLTLSMLTDPSHITAEAKRL
ARGSPSLASASQLSAPSLKATCTTHHSDPSDDALIEANLLWRQEMGNITRVESN
KVWILDSFELHAGDEBRESIAABEILIRKSKFPSPALPIWARPDNPPLLESWKDPY
VPPVHVCGLPPTKAPPIPPRRKRTTVLTESNVSALELATKTFSGSGSVAVDSST
ATAPLDASDDGDKGVESHGSMPLLEGPDPDLSDGSMSTVSEASBDSVCCMS
YTWGALI.TPCAARESKLPINLSNLLRHHNVYATTSRASLQRKQVTFDRLQVID
DHRVDULKENKAKASTVKALLSIEBACKLTTPHASAKSKFYGKADKVRNLSRAVNH
RSWMDLEDETETPIDITIMAKSEVFCQPEKGRKPARLIVFPDLGVRVCEKMALVD
VSTLPQAVNGSSYFGQYSPKQRFVFLNTWKSCKPMGFSDYTRCFDSTVETSDIRV
EESIYQCCDLAPARQAIRSLITERLYIGGPLTNSKGQNCQYRRCRASGVLTTSNGNLT

CDS

FEATURES
source

ORIGIN

Query Match 98.7%; Score 9473.6; DB 13; Length 9596;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 9526; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 1 GCACAGCCCCCTCATGCGGGCGACACTCCACCATGAATCACTCCCTCTGTGAGGAACACTG 60
DB 1 GCACAGCCCCCTCATGCGGGCGACACTCCACCATGAATCACTCCCTCTGTGAGGAACACTG 60
QY 61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTGTCGTGACGCTCCAGGAC 120
DB 61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTGTCGTGACGCTCCAGGAC 120
QY 121 CCCCCCTCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCAG 180
DB 121 CCCCCCTCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCAG 180
QY 181 GACGACCGGCTCCTTTCTTGGATC-AACCGCTCAATGCTCGAGATTTTGGCGGTGCCCC 239
DB 181 GACGACCGGCTCCTTTCTTGGATCAAAACCGCTCAATGCTCGAGATTTGGCGGTGCCCC 240
QY 240 CGCGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTTGTGTACTGTGCTGATAG 299
DB 241 CGCGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGGCTTTGTGTACTGTGCTGATAG 300
QY 300 GGTGCTTGGAGTGCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAGCAATCTCTAAA 359
DB 301 GGTGCTTGGAGTGCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAGCAATCTCTAAA 360
QY 360 CTTCAAGAAAAACCAAAAGTAAACCAACCGCGGCCACGAGCAGTCAAGTTTCCCGGGC 419
DB 361 CTTCAAGAAAAACCAAAAGTAAACCAACCGCGGCCACGAGCAGTCAAGTTTCCCGGGC 420
QY 420 GGTGCTCAGATCGTGGTGAGTTTACTGTGTCGCGGCGAGGGGCCCCAGGTTGGGTGTG 479
DB 421 GGTGCTCAGATCGTGGTGAGTTTACTGTGTCGCGGCGAGGGGCCCCAGGTTGGGTGTG 480
QY 480 CGCGGACTAGGAAGGTTCCGAGCGGTGCGAACCTCTGTGGAAGGCGACAACTATCCCA 539
DB 481 CGCGGACTAGGAAGGTTCCGAGCGGTGCGAACCTCTGTGGAAGGCGACAACTATCCCA 540
QY 540 AAGGCTCGCGACCGAGGCGAGGCTGGGCTCAGCCCGGGTACCTTGGCCCTCTAT 599
DB 541 AAGGCTCGCGACCGAGGCGAGGCTGGGCTCAGCCCGGGTACCTTGGCCCTCTAT 600
QY 600 GGCAATGAGGGCTTGGGGTGGGAGAGTGGCTCTGTGACCCCGGGCTCCCGGCTTAGT 659
DB 601 GGCAATGAGGGCTTGGGGTGGGAGAGTGGCTCTGTGACCCCGGGCTCCCGGCTTAGT 660
QY 660 TGGGGCCCCACGACCCCCCGGCTAGGTTCGCGTAACTTTGGGTAAAGTTCATGATACCTT 719
DB 661 TGGGGCCCCACGACCCCCCGGCTAGGTTCGCGTAACTTTGGGTAAAGTTCATGATACCTT 720
QY 720 ACATGGGCTTCGCCGATCTCTGAGGGTATATTCGGCTCGCTCGGCGCCCCCTAGGGGGC 779
DB 721 ACATGGGCTTCGCCGATCTCTGAGGGTATATTCGGCTCGCTCGGCGCCCCCTAGGGGGC 780
QY 780 GCTGCCAGGGCTTGGCACACGGTGTCCGGGTTCTGAGGAGCGCGGTGAACATATGCAACA 839
DB 781 GCTGCCAGGGCTTGGCACACGGTGTCCGGGTTCTGAGGAGCGCGGTGAACATATGCAACA 840
QY 840 GGGAACTTGGCCGGTTCCTTCTTCTATCTTCTTCTTGGCTGTGCTGTCTGTGTTGACC 899
DB 841 GGGAACTTGGCCGGTTCCTTCTTCTATCTTCTTCTTGGCTGTGCTGTCTGTGTTGACC 900
QY 900 ATCCAGCTTCCGCTTATGAAGTCCGCAACGTGTCCGGGATATACCATGTACAGACGAC 959

901 ATCCAGCTTCGGCTTATGAAGTGGCAACGTGTCGGGATATACCATGTCAACGACGAC 960
960 TGCTCCAACTCAAGCAATTTGTATGAGCGAGGAGCGATGATCATGCTACCTCCCGGTGC 1019
961 TGCTCCAACTCAAGCAATTTGTATGAGCGAGGAGCGATGATCATGCTACCTCCCGGTGC 1020
1020 GTGCCCTGTGTTGAGGAGGTAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCCAAGCTC 1079
1021 GTGCCCTGTGTTGAGGAGGTAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCCAAGCTC 1080
1080 GCGGCAGGAATGCCAGCGTCCCACTACGACAATACGACGCAAGCTCGCATTTGCTGTT 1139
1081 GCGGCAGGAATGCCAGCGTCCCACTACGACAATACGACGCAAGCTCGCATTTGCTGTT 1140
1140 GGGACGGCTGTTCTGCTCCGCTATGATCGTGGGGATCTCTGCGGATCTATTTTCCTC 1199
1141 GGGACGGCTGTTCTGCTCCGCTATGATCGTGGGGATCTCTGCGGATCTATTTTCCTC 1200
1200 GTCTCCAGCTGTTCACTTCTCGCTCCCGGATGAGACAGTGCAGACTGCNACTGC 1259
1201 GTCTCCAGCTGTTCACTTCTCGCTCCCGGATGAGACAGTGCAGACTGCNACTGC 1260
1260 TCAATCTATCCCGGCATGATACAGTCAACGATGGCTTGGGATGATGATGAATCTGG 1319
1261 TCAATCTATCCCGGCATGATACAGTCAACGATGGCTTGGGATGATGATGAATCTGG 1320
1320 TCACCTACAACAGCCCTAGTGTGTCGAGTTGCTCCGGATCCCAAGCTGTGCTGGAC 1379
1321 TCACCTACAACAGCCCTAGTGTGTCGAGTTGCTCCGGATCCCAAGCTGTGCTGGAC 1380
1380 ATGGTGGGGGGCCCACTGGGAGTCTTGGGGGCTTGGCTACTATTCATGCTAGGG 1439
1381 ATGGTGGGGGGCCCACTGGGAGTCTTGGGGGCTTGGCTACTATTCATGCTAGGG 1440
1440 AACTGGGCTTAAGTTCTGATGTGCGCTACTCTTTGCGGGGTTGAGCGGAGACCCAC 1499
1441 AACTGGGCTTAAGTTCTGATGTGCGCTACTCTTTGCGGGGTTGAGCGGAGACCCAC 1500
1500 ACGAGGGGAGGTGGCGGCACACCACTCCGGTTCACTCCGTTTCACTCATCTGG 1559
1501 ACGAGGGGAGGTGGTGGCGGCACCACTCCGGTTTACGTCCTTTCTCATCTGG 1560
1560 GGTCTCAGAAATCCAGTTGTGATATCAACAGGAGCTGCGACATCAAGGACTGCG 1619
1561 GGTCTCAGAAATCCAGTTGTGATATCAACAGGAGCTGCGACATCAAGGAGTCTGCT 1620
1620 CTAAATTCGAATGACTCCCTCCAACTGGGTTCTTTGCGCGCTGTTTTCAGCACACAAG 1679
1621 CTAAATTCGAATGACTCCCTCCAGACTGGGTTCTTTGCGCGCTGTTTTCAGCACACAAG 1680
1680 TTCAACTCTCGGGTGCCCGAGGCGATGGCAGCTGCGGCCCATTTGACTGTTTCGCC 1739
1681 TTCAACTCTCGGGTGCCCGAGGCGATGGCAGCTGCGGCCCATTTGACTGTTTCGCC 1740
1740 CAGGGGTGGGGCCCATCACTATACCTAAGCTTAACAGCTCGGATCAGAGGCTTATTGC 1799
1741 CAGGGGTGGGGCCCATCACTATACCTAAGCTTAACAGCTCGGATCAGAGGCTTATTGC 1800
1800 TGGCATTAACGCGCTCGACCGTGTGTTGTCGTACCGCGCTCGAGGTGTCGTCCAGTG 1859
1801 TGGCATTAACGCGCTCGACCGTGTGTTGTCGTACCGCGCTCGAGGTGTCGTCCAGTG 1860
1860 TATTGTTTCAACCCAGCGCTGTTGTTGGGACCAACGATCGTTTCCGGTGTCCCTAG 1919
1861 TATTGTTTCAACCCAGCGCTGTTGTTGGGACCAACGATCGTTTCCGGTGTCCCTAG 1920
1920 TATAGCTGGGGGAGAAATGAGACAGAGTGTGCTCCTCAACAAACGCGTCCGCCACAA 1979
1921 TATAGCTGGGGGAGAAATGAGACAGAGTGTGCTCCTCAACAAACGCGTCCGCCACAA 1980
1980 GGCACCTGGTTCGGCTGATCATGGATGAATAGTACCTGGGTTCACTAAGACGTGGGAGT 2039

1981 GGCACCTGGTTCGGCTGATCATGGATGATAGTACTGGGTTCACTAAGACGTGGGAGGC 2040
2040 CCCCGGTGAACATCGGGGGGTGGTAACGACACTTTGATCTCCGCCACGACTGCTTC 2099
2041 CCCCGGTGAACATCGGGGGGTGGTAACGACACTTTGATCTCCGCCACGACTGCTTC 2100
2100 CGGAAGCACCCCGAGGTACTTACACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGG 2159
2101 CGGAAGCACCCCGAGGTACTTACACAAAATGTGGCTCGGGGCCCTGGTTGACACCTAGG 2160
2160 TGCTAGTAGACTACCATACAGGCTTTGGCACTACCCCTGCGCATCTCAATTTTTTCATC 2219
2161 TGCTAGTAGACTACCATACAGGCTTTGGCACTACCCCTGCGCATCTCAATTTTTTCATC 2220
2220 TTTAAGGTTAGGATGTATGTGGGGGGGTGGAGCACAGGCTCAATGGCGCATGAATGG 2279
2221 TTTAAGGTTAGGATGTATGTGGGGGGGTGGAGCACAGGCTCAATGGCGCATGAATGG 2280
2280 ACTCGAGGAGCGCTGTAACTTGGAGACAGGATAGGTGAGAACTCAGCCCGCTGCTG 2339
2281 ACTCGAGGAGCGCTGTAACTTGGAGACAGGATAGGTGAGAACTCAGCCCGCTGCTG 2340
2340 CTGTCTAACACAGAGTGGCAGATCTGCCCTGTGCCCTTCAACCCCTACCGGCTTTATCC 2400
2400 ACTGGTTTGATCCATCTCCATCAGAACATCGTGGACGTGCAATACCTGTAGCGGTAGGG 2459
2401 ACTGGTTTGATCCATCTCCATCAGAACATCGTGGACGTGCAATACCTGTAGCGGTAGGG 2460
2460 TCAGCTTTGTCTCTTTTGGCAATCAAATGGAGTACATCTCTGTGTCTTTCTCTCTCTG 2519
2461 TCAGCTTTGTCTCTTTTGGCAATCAAATGGAGTACATCTCTGTGTCTTTCTCTCTCTG 2520
2520 GCAGACGCGCGTGTGTGCTCTGTGGATGATGCTGTGATGACCTAGCCAGGCTGAGGCC 2579
2521 GCAGACGCGCGTGTGTGCTCTGTGGATGATGCTGTGATGACCTAGCCAGGCTGAGGCC 2580
2580 GCCTTAGAGAACTTGTGGTCTCTCAATCGCGCGTCCGTGGCGGAGCGCATGTTATCTC 2639
2581 GCCTTAGAGAACTTGTGGTCTCTCAATCGCGCGTCCGTGGCGGAGCGCATGTTATCTC 2640
2640 TCCTTTCTGTGTTCTCTCGCGCTGTACATTAAGGCGAGCTCGCTCCTCGGGCG 2699
2641 TCCTTTCTGTGTTCTCTCGCGCTGTGTGATGATGCTGTGATGAGCGCTGCTCCTCGGGCG 2700
2700 GCGTATGCTTTTATGGCGTATGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2759
2701 GCGTATGCTTTTATGGCGTATGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
2760 GCTTACGCTTTGACCGGGAGATGGCTGCATCGTGGGGGTGGCGGTTCTGTAGGTTCTG 2819
2761 GCTTACGCTTTGACCGGGAGATGGCTGCATCGTGGGGGTGGCGGTTCTGTAGGTTCTG 2820
2820 GTATTCTTGACTGTGCTACCATCTACAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2879
2821 GTATTCTTGACTGTGCTACCATCTACAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2880
2880 CAATACTTTATCACCAGAGCGGAGCGACATGCAAGTGTGGGTCCCCCCCCCTCAACGTT 2939
2881 CAATACTTTATCACCAGAGCGGAGCGACATGCAAGTGTGGGTCCCCCCCCCTCAACGTT 2940
2940 GGGGAGGCGCGATGCCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2999
2941 GGGGAGGCGCGATGCCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3000
3000 GACATCAACAACTCCTCTCGCCATCTCGGCCCGCTCATGCTGCTCAGGCTGGCAT 3059
3001 GACATCAACAACTCCTCTCGCCATCTCGGCCCGCTCATGCTGCTCAGGCTGGCAT 3060
3060 ACGAGAGTGGCGTCTCTCTGTCGCGCTCAAGGCTCATTCGTGCATGCTGTTAGTGGCA 3119
3061 ACGAGAGTGGCGTCTCTCTGTCGCGCTCAAGGCTCATTCGTGCATGCTGTTAGTGGCA 3120

QY 3120 AAAGTCGCGGGGTCAATATGTCAAATGGTCTTATGAAGCTGGGCGGCTGACAGGT 3179
DB |||||
QY 3121 AAAGTCGCGGGGTCAATATGTCAAATGGCTTCAATGAAGCTGGGCGGCTGACAGGT 3180
DB |||||
QY 3180 AGTACGTTTATAACCATCTTACCCCACTCGGGGACTGGGCCCAACGCGGGCTTACGAGAC 3239
DB |||||
QY 3181 AGTACGTTTATAACCATCTTACCCCACTCGGGGACTGGGCCCAACGCGGGCTTACGAGAC 3240
DB |||||
QY 3240 CTTGCGGTGGCGTAGAGCCGCTGCTCTTCTCGCCCATGGAGACCAAGGTCAATCACTGG 3299
DB |||||
QY 3241 CTTGCGGTGGCGTAGAGCCGCTGCTCTTCTCGCCCATGGAGACCAAGGTCAATCACTGG 3300
DB |||||
QY 3300 GGAGCAGACACCCGCTGGTGGGACATCATCTTGGGTCTACCCGTCTCCGCCCAAGG 3359
DB |||||
QY 3301 GGAGCAGACACCCGCTGGTGGGACATCATCTTGGGTCTACCCGTCTCCGCCCAAGG 3360
DB |||||
QY 3360 GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTTGGCGACTCCCT 3419
DB |||||
QY 3361 GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTTGGCGACTCCCT 3420
DB |||||
QY 3420 GCGCCCATACGCGCTACTCCCAACAAACGCGGGGCTACTTGGTTGCATCATCACTAGC 3479
DB |||||
QY 3421 GCGCCCATACGCGCTACTCCCAACAAACGCGGGGCTACTTGGTTGGCTCATCACTAGC 3480
DB |||||
QY 3480 CTCACAGGCGGGACAAAGAACCAAGGTCGAAGGGAGGTTCAAGTGGTTCTACCGCAACA 3539
DB |||||
QY 3481 CTCACAGGCGGGACAAAGAACCAAGGTCGAAGGGAGGTTCAAGTGGTTCTACCGCAACA 3540
DB |||||
QY 3540 CAATCTTCTCGGGACCTGCATCAACGGGTGTGTGGACTGTCTACCATGGCGTGGC 3599
DB |||||
QY 3541 CAATCTTCTCGGGACCTGCATCAACGGGTGTGTGGACTGTCTACCATGGCGTGGC 3600
DB |||||
QY 3600 TCGAAGACCTTAGCGGTCCAAAGGTCCAATCAACCCAAATGTACACCAATGTAGACCTG 3659
DB |||||
QY 3601 TCGAAGACCTTAGCGGTCCAAAGGTCCAATCAACCCAAATGTACACCAATGTAGACCTG 3660
DB |||||
QY 3660 GACCTCGTGGGTGGCAGGCGCCCCCGGGCGGCTCCATGACACCATGACAGCTGTGGC 3719
DB |||||
QY 3720 AGCTCGGACCTTACTTTGGTCAAGACATGCTGATGTCAATCCGTTGGCGCGGAGGC 3779
DB |||||
QY 3721 AGCTCGGACCTTACTTTGGTCAAGACATGCTGATGTCAATCCCGGTGGCGCGGAGGC 3780
DB |||||
QY 3780 GACAGAGGGGAGTCTACTCTCCCGCAGCGCGCTCTCTACCTGAAGGCTCTCTCGGGT 3839
DB |||||
QY 3781 GACAGAGGGGAGTCTACTCTCCCGCAGCGCGCTCTCTACCTGAAGGCTCTCTCGGGT 3840
DB |||||
QY 3840 GGTCCATTGCTTTGCCCTTCGGGGCAGCTGTGGGGCGTCTTCGGGGCTGTGTGCAAC 3899
DB |||||
QY 3841 GGTCTTTGCTTTGCCCTTCGGGGCAGCTGTGGGGCGTCTTCGGGGCTGTGTGCAAC 3900
DB |||||
QY 3900 CGGGGGTTCGGAAGGCGGTGGAATTCATACCCGTTGAGTCTATGGAATACCATGCGG 3959
DB |||||
QY 3901 CGGGGGTTCGGAAGGCGGTGGAATTCATACCCGTTGAGTCTATGGAATACCATGCGG 3960
DB |||||
QY 3960 TCTCGGTCTTCAAGACAACTCAACCCCGGCTGTACCGCAGACATTCCTCAAGTGGA 4019
DB |||||
QY 3961 TCTCGGTCTTCAAGACAACTCAACCCCGGCTGTACCGCAGACATTCCTCAAGTGGA 4020
DB |||||
QY 4020 CATCTGACGCTCTCTACTGGCAGCGGCAAGAGCACCAAGTGCAGGCTGTGTATGACGCC 4079
DB |||||
QY 4021 CATCTGACGCTCTCTACTGGCAGCGGCAAGAGCACCAAGTGCAGGCTGTGTATGACGCC 4080
DB |||||
QY 4080 CAAGGGTACAAGGTGCTGCTCTGAACCCGCTCGGTGCGGCACTTATAGGTTTGGGGCG 4139
DB |||||
QY 4081 CAAGGGTACAAGGTGCTGCTCTGAATCCGTTCCGTTGCGGCACTTATAGGTTTGGGGCG 4140
DB |||||
QY 4140 TATATGTCGAAGGCACGCGTATGACCCCTTAACATCAGAACTGGGGTGAAGCACTTACC 4199
DB |||||
QY 4141 TATATGTCGAAGGCACGCGTATGACCCCTTAACATCAGAACTGGGGTGAAGCACTTACC 4200
DB |||||

QY 4200 ACGGGCGGCTCCATTACGTACTCCACCTATGCAAGTTCTTTGCCGACGGTGGCTGTCT 4259
DB |||||
QY 4201 ACGGGCGGCTCCATTACGTACTCCACCTATGCAAGTTCTTTGCCGACGGTGGCTGTCT 4260
DB |||||
QY 4260 GGGGGCGGCTATGACATCATATAATATGTGATGTCACATCAACTGACTCGACTACATC 4319
DB |||||
QY 4261 GGGGGCGGCTATGACATCATATAATATGTGATGTCACATCAACTGACTCGACTACATC 4320
DB |||||
QY 4320 TTGGGCATCGGCACAGTCTGGAACCAAGCGGAGACGGCTGGAGCGGGCTCGTCTGCTC 4379
DB |||||
QY 4321 TTGGGCATTTGGCACAGTCTGGAACCAAGCGGAGACGGCTGGAGCGGGCTTTGCTGCTC 4380
DB |||||
QY 4380 GGCACCGCTACACCTCCGGGATCGGTACCGTGGCCACACCCCAATATCGAGGAAATAGGC 4439
DB |||||
QY 4381 GGCACCGCTACACCTCCGGGATCGGTACCGTGGCCACACCCCAATATCGAGGAAATAGGC 4440
DB |||||
QY 4440 CTGTCCAAACAATGGAGATCCCTTTATGGAAGCAAGCATCCCAATGAGGGCCATCAAG 4499
DB |||||
QY 4441 CTGTCCAAACAATGGAGATCCCTTTATGGAAGCAAGCATCCCAATGAGGGCCATCAAG 4500
DB |||||
QY 4500 GGGGGAGGCTCTCATTTTCTGCCATTTCCAGAAGAAATGTGACGAGCTCGCGGCAAG 4559
DB |||||
QY 4501 GGGGGAGGCTCTCATTTTCTGCCATTTCCAGAAGAAATGTGACGAGCTCGCGGCAAG 4560
DB |||||
QY 4560 CTGACAGGCTCGGACTGAAAGCTGTAGCATATTTACCGGGCTTCTGATGTCCGTGATA 4619
DB |||||
QY 4561 CTGACAGGCTCGGACTGAAAGCTGTAGCATATTTACCGGGCTTCTGATGTCCGTGATA 4620
DB |||||
QY 4620 CCGCTTATCGGAGACGCTGTTGCTGGCAACAGACGCTCTAATGACGGGTTTACCGGC 4679
DB |||||
QY 4621 CCGCTTATCGGAGACGCTGTTGCTGGCAACAGACGCTCTAATGACGGGTTTACCGGC 4680
DB |||||
QY 4680 GATTTTGACTCAGTGATTCGACTGCAATACATGTGTCACTCAGACAGTTCGACTTG 4739
DB |||||
QY 4681 GATTTTGACTCAGTGATTCGACTGCAATACATGTGTCACTCAGACAGTTCGACTTG 4740
DB |||||
QY 4740 GATCCCACTTCCACTTGAGACGACCGTCCCAAGACGCTGTCGCGTCCGCA 4799
DB |||||
QY 4741 GATCCCACTTCCACTTGAGACGACCGTCCCAAGACGCTGTCGCGTCCGCA 4800
DB |||||
QY 4800 CGCGAGGTAGAACTGGCAGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAA 4859
DB |||||
QY 4801 CGCGAGGTAGAACTGGCAGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAA 4860
DB |||||
QY 4860 CGCCCTCGGGCATGTTTCGATTTTCGCTCTGTGTGAGTGTATGACGCGGGCTGTGCT 4919
DB |||||
QY 4861 CGCCCTCGGGCATGTTTCGATTTTCGCTCTGTGTGAGTGTATGACGCGGGCTGTGCT 4920
DB |||||
QY 4920 TGGTATGAGCTCAGCCCGCTGAGACCTCGGTTAGGTTGGGGCTTACCTAAATACACCA 4979
DB |||||
QY 4921 TGGCATGAGCTCAGCCCGCTGAGACCTCGGTTAGGTTGGGGCTTACCTAAATACACCA 4980
DB |||||
QY 4980 GGGTTGCCCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTCACAGGCTCACCC 5039
DB |||||
QY 4981 GGAATGCCCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTCACAGGCTCACCC 5040
DB |||||
QY 5040 CACATAGATGCCACTTTCCTGTCCAGACTAAACAGGACGAGAGACAACTTTCCTTACCTG 5099
DB |||||
QY 5041 CACATAGATGCCACTTTCCTGTCCAGACTAAACAGGACGAGAGACAACTTTCCTTACCTG 5100
DB |||||
QY 5100 GTGGCATATCAGGTACAGTGTGGCCAGGCTCAAGCTCCACCTCCATCGTGGGACCAA 5159
DB |||||
QY 5101 GTGGCATATCAGGTACAGTGTGGCCAGGCTCAAGCTCCACCTCCATCGTGGGACCAA 5160
DB |||||
QY 5160 ATGTGGAAGTGTCTCATACGCTGAAACCTACACTGCACGGGCCAAACCCCTCTGTAT 5219
DB |||||
QY 5161 ATGTGGAAGTGTCTCATACGCTGAAACCTACACTGCACGGGCCAAACCCCTCTGTAT 5220
DB |||||
QY 5220 AGGTAGGAGCGGTCCTCAAAATGAGGTCTCTCAACACACCCCAATAAATACATCATG 5279
DB |||||
QY 5221 AGGTAGGAGCGGTCCTCAAAATGAGGTCTCTCAACACACCCCAATAAATACATCATG 5280
DB |||||
QY 5280 GCATGATGTGGGTGACCTGGAGGTCGTCTACTAGCACCTGGGTGCTGGTAGGGGAGTC 5339
DB |||||

5281	Db		GCATGCAATGTCGGCTGACCTTGGAGGTCGTCACTAGCACCTTGGGTGCTGGTGGCGGAGTC	5340
5340	Qy	CTTGCAGCTTTGGCCGCATATCTGCTCAACACAGGCAGTGTGTCATTGTGGGCGAGATC	5399	
5341	Db	CTTGCACTTTTGGCCGCATATCTGCTGACGACAGGCAGTGTGTCATTGTGGGCGAGATC	5400	
5400	Qy	ATCTTGTCCGGGAAGCCAGCTGTCTGTTTCCGACAGGGAAGTCTCTACACAGGAGTTTCGAT	5459	
5401	Db	ATCTTGTCCGGGAAGCCAGCTGTCTGTTTCCGACAGGGAAGTCTCTACACAGGAGTTTCGAT	5460	
5460	Qy	GAGATGAAGATGTGGCTCAACAATTCCTTACATCGAGCAGGAATGCACTGCGCGAG	5519	
5461	Db	GAGATGAAGATGTGGCTCAACAATTCCTTACATCGAGCAGGAATGCACTGCGCGAG	5520	
5520	Qy	CAATTCAGGCAAAAGGCGCTCGGCTTGCTGCAACAGCGCCACCAAGCAAGCGGAGGCTGCT	5579	
5521	Db	CAATTTAAGCAGAAGGCGCTCGGATTTGTGAAACGGCCACCAAGCAAGCGGAGGCTGCT	5580	
5580	Qy	GCTCCCGTGTGGAGTCCAAAGTGGCGAGCCCTTTGAGACCTTCTGGGCGGAAGCACATGTGG	5639	
5581	Db	GCTCCCGTGTGGAGTCCAAAGTGGCGAGCCCTCGAGACCTTCTGGGCGGAAGCACATGTGG	5640	
5640	Qy	AATTTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCTGCTGGAACCCCGCG	5699	
5641	Db	AATTTTCATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCTGCTGGAACCCCGCG	5700	
5700	Qy	ATAGCATCATTGATGGCATTTACAGCTTCTATCACTAGCCGCTCAACCCCAAAACACC	5759	
5701	Db	ATAGCATCATTGATGGCATTTACAGCTTCTATCACTAGCCGCTCAACCCCAAAACACC	5760	
5760	Qy	CTCCTGTTTAAACATCTTTGGGGGATGGTGGCTGCCCAAATCGCTCTCTCCAGCGTGGC	5819	
5761	Db	CTCCTGTTTAAACATCTTTGGGGGATGGTGGCTGCCCAAATCGCTCTCTCCAGCGTGGC	5820	
5820	Qy	TCAGCTTTCTGTGGCGCGCGCATCGCCGAGCGGCTGTTGGCAGCATAGGCCTTGGGAAG	5879	
5821	Db	TCAGCTTTCTGTGGCGCGCGCATCGCCGAGCGGCTGTTGGCAGCATAGGCCTTGGGAAG	5880	
5880	Qy	GTGCTCGTGGACATCTTGGCGGGCTATGGGCGAGGGTAGCCGCGGCACCTCGTGGCGCTTT	5939	
5881	Db	GTGCTCGTGGACATCTTGGCGGGCTATGGGCGAGGGTAGCCGCGGCACCTCGTGGCGCTTT	5940	
5940	Qy	AAGGTCATGAGCGGCGAGTGCCCTCCACCGAGGACCTGGTGCAATTTACTTCCCTGCCATC	5999	
5941	Db	AAGGTCATGAGCGGCGAGTGCCCTCCACCGAGGACCTGGTGCAATTTACTTCCCTGCCATC	6000	
6000	Qy	CTCTCTCTGTGCCCTGGTCGTGGGGTGTGTGCGCAGCAATATCGCTGCGCACGTG	6059	
6001	Db	CTCTCTCTGTGCCCTGGTCGTGGGGTGTGTGCGCAGCAATATCGCTGCGCACGTG	6060	
6060	Qy	GGCCCGGAGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGGGT	6119	
6061	Db	GGCCCGGAGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGGGT	6120	
6120	Qy	AACACAGTCTCCCTTACGCATCTATGTGCTCAGAGCGACGCTGCAGACAGTGTCACTCAG	6179	
6121	Db	AACACAGTCTCCCTTACGCATCTATGTGCTCAGAGCGACGCTGCAGACAGTGTCACTCAG	6180	
6180	Qy	ATCCTTCTTAGCCCTTACCATCACTCAATGCTGAAGCGGCTCCACCAGTGGATTAATGAG	6239	
6181	Db	ATCCTTCTTAGCCCTTACCATCACTCAATGCTGAAGCGGCTCCACCAGTGGATTAATGAG	6240	
6240	Qy	GACTGCTCTAGCCCATGCTCGGCTCTGGGCTTAAGGATGTTTGGGATTTGGATATGCAAG	6299	
6241	Db	GACTGCTCTAGCCCATGCTCGGCTCTGGGCTTAAGGATGTTTGGGATTTGGATATGCAAG	6300	
6300	Qy	GTGTTGACTGACTTCAAGACCTGGCTCCAGTCCAAACCTCTCTGCCGCGGTTACCGGGATC	6359	
6301	Db	GTGTTGACTGACTTCAAGACCTGGCTCCAGTCCAAACCTCTCTGCCGCGGTTACCGGGATC	6360	
6360	Qy	CTTTTCTGTGATGCCAACCGGGGTACAAGGAGTCTTGGCGGGGGGACGGCATCATGCAA	6419	

Db	6361	CTTTTCTGTCTATGCGCAACCGGGGTACAAGGGAGTCTGGCGGGGGNACGCATCATGCAA	6420
Qy	6420	ACCACCTGCCCATCGGAGACACAGATCGCGGACATGTCAAAAACGGTTCATGAGGATC	6479
Db	6421	ACCACCTGCCCATGTGGAGCACAGATCGCGGACATGTCAAAAACGGTTCATGAGGATC	6480
Qy	6480	GTAGGGCTAGAACTCTGACGAACAAGTGGAACGGAAGTTCGCCATCAACGCATACACC	6539
Db	6481	GTAGGGCTTAGGACCTGACGAACAAGTGGAACGGAAGTTCGCCATCAACGCATACACC	6540
Qy	6540	ACGGGACCTTGCACACCCCTCCCGGGCCCAACTATTCCAGGGCGCTATGCGGGTGGCT	6599
Db	6541	ACGGGACCTTGCACACCTTCCCGGGCCCAACTATTCCAGGGCGCTATGCGGGTGGCT	6600
Qy	6600	GCTGAGGAGTACGTGGAGGTACGCGTGTGGGGATTTCCACTACGTGACGGGCATGACC	6659
Db	6601	GCTGAGGAGTACGTGGAGGTACGCGTGTGGGGATTTCCACTACGTGACGGGCATGACC	6660
Qy	6660	ACTGACAAAGTGTGCCATATGCAAGTTCGGGCCCCGAAATTTTTCACGAGGTGGAT	6719
Db	6661	ACTGACAAAGTGTGCCATATGCAAGTTCGGGCCCCGAAATTTTTCACGAGGTGGAT	6720
Qy	6720	GGAGTGGGTTGCACAGGTACGCTCGGGCTGCAGACCTCTCTACGGGAGGACGTCAAG	6779
Db	6721	GGAGTGGGTTGCACAGGTACGCTCGGGCTGCAGACCTCTCTACGGGAGGACGTCAAG	6780
Qy	6780	TTCCAGGTCCGGCTCAACCAATATTCTGGTCCGGTCGAGCTCCCATCGAGCCCCGAACCG	6839
Db	6781	TTCCAGGTCCGGCTCAACCAATATTCTGGTCCGGTCGAGCTCCCATCGAGCCCCGAACCG	6840
Qy	6840	GACGTAAACAGTGTACTTCCATGCTCACCGATCCCTCCCAATTCACGACGAGACGGCT	6899
Db	6841	GACGTAAACAGTGTACTTCCATGCTCACCGATCCCTCCCAATTCACGACGAGACGGCT	6900
Qy	6900	AAGGTAGGCTGGCTAGAGGTTCCCGGCTCTTTAGCCAGCTCATCAGCTAGCCAGTTG	6959
Db	6901	AAGGTAGGCTGGCTAGAGGTTCTCCCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTTG	6960
Qy	6960	TCTCGGCTCTTTTGAAGGCGACATGACCTACCCACCATGACTCCCCGGACGCTGACCTC	7019
Db	6961	TCTCGGCTCTTTTGAAGGCGACATGACCTACCCACCATGACTCCCCGGACGCTGACCTC	7020
Qy	7020	ATCAGAGCCAACTCTTTGTGGCGCAGGAGATGGCGGAAACATCACTCGCGTGGAGTCA	7079
Db	7021	ATCAGAGCCAACTCTTTGTGGCGCAGGAGATGGCGGAAACATCACTCGCGTGGAGTCA	7080
Qy	7080	GAGAAATAGGTAGTAAATCTTGACTCTTTCGAACCGCTTCAACGGGGGGGATGAGGG	7139
Db	7081	GAGAAATAGGTAGTAAATCTTGACTCTTTCGAACCGCTTCAACGGGGGGGATGAGGG	7140
Qy	7140	GAGATATCCGTTCGGCGGAGATCCTTCGCAAAATTCAGGAAGTTCCCTCAGCGCTGCCC	7199
Db	7141	GAGATATCCGTTCGGCGGAGATCCTTCGCAAAATTCAGGAAGTTCCCTCAGCGCTGCCC	7200
Qy	7200	ATATGGGACGCCCCGACTCAATTCCTCACTGCTAGAGTCTCTGGAAGGACCCGGAATAC	7259
Db	7201	ATATGGGACGCCCCGACTCAATTCCTCACTGCTAGAGTCTCTGGAAGGACCCGGAATAC	7260
Qy	7260	GTCCCTCGGGTGTACAGGATGCCATTTGCCACTACCAAGGCTCTCTCAATACCACCT	7319
Db	7261	GTCCCTCGGGTGTACAGGATGCCATTTGCCACTACCAAGGCTCTCTCAATACCACCT	7320
Qy	7320	CCAGGAGAAAGAGACGGTTGTCTGCACAGAATCCAAATGTCTTCTGCCCTGGCGGAG	7379
Db	7321	CCAGGAGAAAGAGACGGTTGTCTGCACAGAATCCAAATGTCTTCTGCCCTGGCGGAG	7380
Qy	7380	CTCGCCACTAAGACTTCCTGGTAGTCTCCGGATCTGTCGGCCGTTGTATAGCGGCACCGGAC	7439
Db	7381	CTCGCCACTAAGACTTCCTGGTAGTCTCCGGATCTGTCGGCCGTTGTATAGCGGCACCGGAC	7440
Qy	7440	GCCCTTCTGTACTTGGCTCCGACGACGGTGAACAAAGATCCGACGTTGAGTCGTACTCC	7499
Db	7441	GCCCTTCTGTACTTGGCTCCGACGACGGTGAACAAAGATCCGACGTTGAGTCGTACTCC	7500

QY	7500	TCCATGCCCCCTTTGAAGGGAGCGCGGGACCCCGATCTCTCAGGGACGGGTCTTGGTCT	7559
DB	7501	TCCATGCCCCCTTTGAAGGGAGCGCGGGACCCCGATCTCAGCGAGGGTCTTGGTCT	7560
QY	7560	ACCGTGATGAGAGGGCTAGTAGAGATGTCTGTCTCAATGTCTATACGTGGACA	7619
DB	7561	ACCGTGATGAGAGGGCTAGTAGAGACGTCTGTCTCAATGTCTATACGTGGACA	7620
QY	7620	GGCGCCCTGATCAGCCATCGCTGCGGAGGAAGTAAAGCTGCCATCAACCCTGTGAGC	7679
DB	7621	GGCGCCCTGATCAGCCATCGCTGCGGAGGAAGTAAAGCTGCCATCAACCCTGTGAGC	7680
QY	7680	AACTCTTTGCTGGGTCAACCAACATGTGTTACGCCACAACATCCCGCAGCGCAAGCCTC	7739
DB	7681	AACTCTTTGCTGGGTCAACCAACATGTGTTACGCCACAACATCCCGCAGCGCAAGCCTC	7740
QY	7740	CGGCAGAAAGGTACCTTTTGAAGATTGCCAAGTCTCTGGATGATCATTAACGGGACGTGA	7799
DB	7741	CGGCAGAAAGGTACCTTTTGAAGATTGCCAAGTCTCTGGATGATCATTAACGGGACGTGA	7800
QY	7800	CTCAAGAGATGAAGCGRAAGCGTCCACAGTTAAAGCTAAAGTCTATCTATAGAGGAG	7859
DB	7801	CTCAAGAGATGAAGCGRAAGCGTCCACAGTTAAAGCTAAAGTCTATCTATAGAGGAG	7860
QY	7860	GCCTGCAAGCTGACGCCGCCACATTCGGGCCAAATCCAAATTTGGCTATGGGCAAAAGAC	7919
DB	7861	GCCTGCAAGCTGACGCCGCCACATTCGGGCCAAATCCAAATTTGGCTATGGGCAAAAGAC	7920
QY	7920	GTCCGGAACCTATCAGCAGGGCGTTAAACCAATCGCTCTGTGTGGAGGACTTGTCTG	7979
DB	7921	GTCCGGAACCTATCAGCAGGGCGTTAAACCAATCGCTCTGTGTGGAGGACTTGTCTG	7980
QY	7980	GAAGACACTGAAACACCAATTTGACACACCATCATGGCAAAAAGTGAGTTTCTCGCTC	8039
DB	7981	GAAGACACTGAAACACCAATTTGACACACCATCATGGCAAAAAGTGAGTTTCTCGCTC	8040
QY	8040	CAACGAGAGGGAGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGACCTGGGAGTT	8099
DB	8041	CAACGAGAGGGAGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGACCTGGGAGTT	8100
QY	8100	CGTGTATGCGAAGAGATGGCCCTTTACGAGGTGTCTCCACCTTCCTCAGGCGGTGATG	8159
DB	8101	CGTGTATGCGAAGAGATGGCCCTTTACGAGGTGTCTCCACCTTCCTCAGGCGGTGATG	8160
QY	8160	GGCTCCTCATACGATTTCAATATCTCCCCCAAGCAGCGGGTCGAGTTCTCTGGTGAATACC	8219
DB	8161	GGCTCCTCATACGATTTCAATATCTCCCCCAAGCAGCGGGTCGAGTTCTCTGGTGAATACC	8220
QY	8220	TGGAAATCAAGAAATGCCCTTATCGGCTTCTCATATGACACCGCTGTTTGTGACTCAACG	8279
DB	8221	TGGAAATCAAGAAATGCCCTTATCGGCTTCTCATATGACACCTGCTGTTTGTGACTCAACG	8280
QY	8280	GTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACAAATGTTGTGACTTGGCCCCC	8339
DB	8281	GTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATTTACAAATGTTGTGACTTGGCCCCC	8340
QY	8340	GAGGCCAGACAGGCCATTAAGGTCTGCTCAAGAGCGGCTTTTACATCGGGGGTCCCTGACT	8399
DB	8341	GAGGCCAGACAGGCCATTAAGGTCTGCTCAAGAGCGGCTTTTACATCGGGGGTCCCTGACT	8400
QY	8400	AACTCAAAAGGGCAGAACTGGGTTATCGCGGTGCCCGCAAGTGGCGGTCTGTGACGACT	8459
DB	8401	AACTCAAAAGGGCAGAACTGGGTTATCGCGGTGCCCGCAAGTGGCGGTCTGTGACGACT	8460
QY	8460	AGCTCGGTTATACCTCACATGTTACTTTGAAGCCACTGCGAGCCTGTGCGAGCTGCAAG	8519
DB	8461	AGCTCGGTTATACCTCACATGTTACTTTGAAGCCACTGCGGCGCTGTGCGAGCTGCAAG	8520
QY	8520	CTCCAGGACTGACCATGCTCTGTGAACGGAGACGACCTTGTCTGTAAGAGCGCG	8579
DB	8521	CTCCAGGACTGACCATGCTCTGTGAACGGAGACGACCTTGTCTGTAAGAGCGCG	8580

RESULT 6
AF054250
LOCUS
DEFINITION

QY 3014 CTTGCTCGCCATATCGGCCCGCTCATGTGCTCCAGGCTGGCATAAACGAGTGCCTGA 3073
Db
QY 3004 CTTGCTCGCCATATCGGCCCGCTCATGTGCTCCAGGCTGGCATAAACGAGTGCCTGA 3063
Db
QY 3074 CTTGCTCGCGCTCAAGGGCTCATTCGTGCATGCAATGTTAGTGCAGAAAAGTGC CGGGG 3133
Db
QY 3064 CTTGCTCGCGCTCAAGGGCTCATTCGTGCATGCAATGTTAGTGCAGAAAAGTGC CGGGG 3123
QY 3134 TCATTATGTCNAATGGTCTTCATGAAGCTGGGCGCTGACAGGTACGTACGTTTATAA 3193
Db
QY 3124 TCATTATGTCNAATGGGCTTCATGAAGCTGGGCGCTGACAGGTACGTACGTTTATAA 3183
QY 3194 CCATCTTACCCACATCGCGGACTGGGCCACGCGGGCTTACGAGACCTTCGGTGGCGGT 3253
Db
QY 3184 CCATCTTACCTGCGGACTGGGCCACGCGGGCTTACGAGACCTTCGGTGGCGGT 3243
QY 3254 AGAGCCCGTCTCTTCGCGCATGGAGACCAAGGTATCATCTGGGGAGCAGACACCGC 3313
Db
QY 3244 AGAGCCCGTCTCTTCGCMCATGGAGACCAAGGTATCATCTGGGGAGCAGACACCGC 3303
QY 3314 TCGGTGCGGACATCATCTTGGGTCTACCGTCTCCGCGCGAAGGGGGAAGGATATT 3373
Db
QY 3304 WSCGTGTGGGACATCATCTTGGGTCTACCGTCTCCGCGCGAAGGGGGAAGGATATT 3363
QY 3374 TTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTGGCGACTCTTCGCGCCCATCACGGC 3433
Db
QY 3364 TTTGGGACCGGCTGATAGTCTCGAAGGGCAAGGTGGCGACTCTTCGCGCCCATCACGGC 3423
QY 3434 CTACTCCCAACAAACGCGGGGCTACTTGGTTGCATCATCACTAGCCCTCACAGGCGGGGA 3493
Db
QY 3424 CTACTCCCAACAAACGCGGGGCTACTTGGTTGCATCATCACTAGCCCTCACAGGCGGGGA 3483
QY 3494 CAAAGAACAGTTCGAAGGGAGGTTCAAGTGTGTTCTACCGCAACAAATCTTCTCTGGC 3553
Db
QY 3484 CAAAGAACAGTTCGAAGGGAGGTTCAAGTGTGTTCTACCGCAACAAATCTTCTCTGGC 3543
QY 3554 GACCTGCATCAACGGGTGTGTGACTGTCTACCATGGCGTGGCTCGAAGACCCTAGC 3613
Db
QY 3544 GACCTGCATCAACGGGTGTGTGACTGTCTACCATGGCGTGGCTCGAAGACCCTAGC 3603
QY 3614 CGGTCCAAAGGTTCCAAATCACCCAAATGTACACCAATGTAGACCTGGACCTCGTGGCTG 3673
Db
QY 3604 CGGTCCAAAGGTTCCAAATCACCCAAATGTACACCAATGTAGACCTGGACCTCGTGGCTG 3663
QY 3674 GCAGGCGCCCCCGGGCGCGCTCCATGACACCATGTGGAGCTCGGACCTTTTA 3733
Db
QY 3664 GCAGGCGCCCCCGGGCGCGCTCCATGACACCATGTGGAGCTCGGACCTTTTA 3723
QY 3734 CTTGGTCAAGACATGCTGATGTCAATCCGGTGGCGCGCGGAGGCGACAGCGGGGAAG 3793
Db
QY 3724 CTTGGTCAAGACATGCTGATGTCAATCCGGTGGCGCGCGGAGGCGACAGCGGGGAAG 3783
QY 3794 TCTACTCTCCCCAGGCGCGCTCTCTACTCTGAAAGGCTCCTCGGGTGGTCCATTGCTTTG 3853
Db
QY 3784 TCTACTCTCCCCAGGCGCGCTCTCTACTCTGAAAGGCTCCTCGGGTGGTCCATTGCTTTG 3843
QY 3854 CCTTCGGGGCAGTGTGGCGGTCTTCGGGCTGTGTGTGACCCCGGGGGGTTCGCGAA 3913
Db
QY 3844 CCTTCGGGGCAGTGTGGCGGTCTTCGGGCTGTGTGTGACCCCGGGGGGTTCGCGAA 3903
QY 3914 GCGGTGGACTTCATACCGTTGAGTCTATGAAACTACCATCGGTCTCCGGTCTTCAC 3973
Db
QY 3904 GCGGTGGACTTCATACCGTTGAGTCTATGAAACTACCATCGGTCTCCGGTCTTCAC 3963
QY 3974 AGACAACTCAACCCCCCGGCTGTACCGCAGACATTCGAAGTGGCAGCATCTCGACGCTCC 4033
Db
QY 3964 AGACAACTCATCCCCCGGCTGTACCGCAGACATTCGAAGTGGCAGCATCTCGACGCTCC 4023
QY 4034 TACTGCGCGGCAAGAGACCAAAAGTGGCGGTGTGTATGACGCCCAAGGGTACAAGT 4093
Db
QY 4024 TACTGCGCGGCAAGAGCACTAAAAGTGGCGGTGTGTATGACGCCCAAGGGTACAAGT 4083

QY 4094 GCTCGTCTGAACCCCGTCCGTTGCCGCCACCTTAGGGTTTGGGCGTATATGTCCAAGGC 4153
Db
QY 4084 GCTCGTCTGAACCCCGTCCGTTGCCGCCACCTTAGGGTTTGGGCGTATATGTCCAAGGC 4143
QY 4154 ACAGGTATCGACCCCTAAACATCAGAACTGGGTAAAGAACCATTTACCACGGCGGCTCCAT 4213
Db
QY 4144 ACAGGTATCGACCCCTAAACATCAGAACTGGGTAAAGAACCATTTACCACGGCGGCTCCAT 4203
QY 4214 TACTACTCTCACCTATGCGCAAGTTCCTTTCGCGCGGTGGTGTCTTCTGGGCGGCTATGA 4273
Db
QY 4204 TACTACTCTCACCTATGCGCAAGTTCCTTTCGCGAVGGTGGTGTCTCYGGGGCGCTATGA 4263
QY 4274 CATCATATATGTGATGAGTGCCACTCAACTGACTCGACTACCATCTTGGGCACTCGGCAC 4333
Db
QY 4264 CATCATATATGTGATGAGTGCCACTCAACTGACTCGACTACCATCTTGGGCACTCGGCAC 4323
QY 4334 AGTCTCGAACAGCGGAGACGGCTGAGCGCGGCTCGTGTGTCTGCGCACCGCTACACC 4393
Db
QY 4324 AGTCTCGAACAGCGGAGACGGCTGAGCGCGGCTTGTGTGTCTGCCACCGCTACACC 4383
QY 4394 TCCGGGATCGGTTACCGTGGCCACACCCCAATATCGAGGAAATAGGCCCTGTCCAAATGG 4453
Db
QY 4384 TCCGGGATCGGTTACCGTGGCCACACCCCAATATCGAGGAAATAGGCCCTGTCCAAATGG 4443
QY 4454 AGAGATCCCGTCTTATGCGCAAGCCATCCCGCATTTAGGCCCATCAAGGGGGAGGCACTCT 4513
Db
QY 4444 AGAGATCCCGTCTTATGCGCAAGCCATCCCGCATTTAGGCCCATCAAGGGGGAGGCACTCT 4503
QY 4514 CATTTCTGCGATTCGAAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCGCTCGG 4573
Db
QY 4504 CATTTCTGCGATTCGAAGAAATGTGACGAGCTCGCGCAAAAGCTGACAGGCGCTCGG 4563
QY 4574 ACTCAACGCTGTAGCATATTACCGGGGCTTGTGTGTCTCGTATACCGCTATTCGGAGA 4633
Db
QY 4564 ACTCAACGCTGTAGCATATTACCGGGGCTTGTGTGTCTCGTATACCGCTATTCGGAGA 4623
QY 4634 CGTCTGTGTCTGTGCAACAGACGCTCTAATGACGGGTTTACCAGCGCATTTTGTACTCAGT 4693
Db
QY 4624 CGTCTGTGTCTGTGCAACAGACGCTCTAATGACGGGTTTACCAGCGCATTTTGTACTCAGT 4683
QY 4694 GATCGACTGCAATACATGTGTCAACAGACAGTGTGAGCTTTCAGCTTGGATTCACCTTCAC 4753
Db
QY 4684 GATCGACTGCAATACATGTGTCAACAGACAGTGTGAGCTTTCAGCTTGGATTCACCTTCAC 4743
QY 4754 CATTTGAGCAGCAGCGCTGCCCCAAGACGGGTGTGCGCTCGCAACGGCGAGGTAGAAC 4813
Db
QY 4744 CATTTGAGCAGCAGCGCTGCCCCAAGACGGGTGTGCGCTCGCAACGGCGAGGTAGAAC 4803
QY 4814 TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAACGGCCCTCGGGCAT 4873
Db
QY 4804 TGGCAGGGGTAGGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAACGGCCCTCGGGCAT 4863
QY 4874 GTTCGATTTCTCGTTCGTGTGATGTATGACCGGGCTGTGCTTGGTATGAGCTCAC 4933
Db
QY 4864 GTTCGATTTCTCGTTCGTGTGATGTATGACCGGGCTGTGCTTGGTATGAGCTCAC 4923
QY 4934 GCGGCTGAGACCTCGGTTAGGTTGCGGCTTACTTAATACACAGGTTTCCCGCTGTG 4993
Db
QY 4924 GCGGCTGAGACCTCGGTTAGGTTGCGGCTTACTTAATACACAGGTTTCCCGCTGTG 4983
QY 4994 CCAGGACCATCTGGAGTTCTGGGAGAGGCTCTTCAACAGGCTCACCCACATAGATGCCCA 5053
Db
QY 4984 CCAGGACCATCTGGAGTTCTGGGAGAGGCTCTTCAACAGGCTCACCCACATAGATGCCCA 5043
QY 5054 CTTCTGTCTCCAGACTAAACAGGAGGAGAGCAACTTTCTCTTACTGTGGGATATCAAGC 5113
Db
QY 5044 CTTCTGTCTCCAGACTAAACAGGAGGAGAGCAACTTTCTCTTACTGTGGGATATCAAGC 5103
QY 5114 TACAGTGTGCGGCGGCTCAAGCTCCACCTCGTGGGACCAAAATGTGGAAGTGTCT 5173
Db
QY 5104 TACAGTGTGCGGCGGCTCAAGCTCCACCTCGTGGGACCAAAATGTGGAAGTGTCT 5163
QY 5174 CATACGGCTGAAACCTACACTGCGCGGCCAACACCCCTGTGTATAGGCTAGGAGCGGT 5233

[illegible]

6244	ATGCTCCGGCTCGTGCTAAGGGATGTTTGGGATTTGGATATGCCACGGTGTGACTGACTT	6303
6314	CAAGACTTGGCTCCAGTCCAAATCTCTGCGGGTTTACCGGGAGTCCCTTCTCTGTCATG	6373
6304	CAAGACTTGGCTCCAGTCCAAAGTCTCTGCGGGTTTACCGGGAGTCCCTTCTCTGTCATG	6363
6374	CCAAACCGGGTCAAGGGAGTCTGGCGGGGGAGCGGCATCATGCAAAACCACTGCCCATG	6433
6364	CCAAACCGGGTCAAGGGAGTCTGGCGGGGGAGCGGCATCATGCAAAACCACTGCCCATG	6423
6434	CGGAGCACAGATCGCCGGACATGTCAAAACGGTTCATAGAGTCTGATGGGCTTGAAC	6493
6424	TGGAGCACAGATCGCCGGACATGTCAAAACGGTTCATAGAGTCTGATGGGCTTGAAC	6483
6494	CTCAGCAACACGTGCGAGAACTTCCCATCAACGCATACACACCGGACCTTGCAC	6553
6484	CTGAGCAACACGTGCGAGAACTTCCCATCAACGCATACACACCGGACCTTGCAC	6543
6554	ACCTTCCCGGGCCCAACTATTTCAGGGCGCTATGGCGGGTGGTGTGAGAGTACGT	6613
6544	ACCTTCCCGGGCCCAACTATTTCAGGGCGCTATGGCGGGTGGTGTGAGAGTACGT	6603
6614	GGAGGTTACCGTGTGGGGATTTCACATGTCGACGGGATGACCACTGACCAACGTA	6673
6604	GGAGGTTACCGTGTGGGGATTTCACATGTCGACGGGATGACCACTGACCAACGTA	6663
6674	GTGCCATGCCAGTTCGGCCCCCGAAATTCCTTCAAGAGGTGATGGAGTGGCGTTGCA	6733
6664	GTGCCATGCCAGTTCGGCCCCCGAAATTCCTTCAAGAGGTGATGGAGTGGCGTTGCA	6723
6734	CAGGTACGCTCCGGCGTGAACCTCTCTACGGGAGGAGCTCACGTTCAGAGTCCGGCT	6793
6724	CAGGTACGCTCCGGCGTGAACCTCTCTACGGGAGGAGCTCACGTTCAGAGTCCGGCT	6783
6794	CAACCAATCTTGGTGGGTTCGAGTCCCATCGAGCCCGAAACCGGACGTAACAGTGCT	6853
6784	CAACCAATCTTGGTGGGTTCGAGTCCCATCGAGCCCGAAACCGGACGTAACAGTGCT	6843
6854	TACTTCCATGCTCACCGATCCCTCCACATTCACAGCAGACCGGCTAAGGCTAGGCTGGC	6913
6844	TACTTCCATGCTCACCGATCCCTCCACATTCACAGCAGACCGGCTAAGGCTAGGCTGGC	6903
6914	TAGAGGTCTCCCGCTCTTTAGCGAGCTCATCAGTACGAGTGTGCTCGGCTCTCTT	6973
6904	TAGAGGTCTCCCGCTCTTTAGCGAGCTCATCAGTACGAGTGTGCTCGGCTCTCTT	6963
6974	GAAGCGACATGCACTACCCACCATGACTCCCGGACGCTGACTCATCGAGGCCAACCT	7033
6964	GAAGCGACATGCACTACCCACCATGACTCCCGGACGCTGACTCATCGAGGCCAACCT	7023
7034	CTTCTGGCGGACAGAGATGGCGGAAACATCAGTTCGGTGGATTCAGAGAAATAGGTAGT	7093
7024	CTTCTGGCGGACAGAGATGGCGGAAACATCAGTTCGGTGGATTCAGAGAAATAGGTAGT	7083
7094	AATTCTGAGCTCTTTCGAAACCGCTTTCACGGGGGGGATGAGAGGAGATATCCGTCGC	7153
7084	AATTCTGAGCTCTTTCGAAACCGCTTTCACGGGGGGGATGAGAGGAGATATCCGTCGC	7143
7154	GGCGGAGATCTCGGAAATCCAGAAAGTTCCTCCTCAGCGTGGCCCATATGGGACGCCC	7213
7144	GGCGGAGATCTCGGAAATCCAGAAAGTTCCTCCTCAGCGTGGCCCATATGGGACGCCC	7203
7214	GGACTACAACTCTCCATCTAGAGTCTGTGAAGGACCGGAGTACGCTCCCTCGGTTGGT	7273
7204	GGACTACAACTCTCCATCTAGAGTCTGTGAAGGACCGGAGTACGCTCCCTCGGTTGGT	7263
7274	ACACGGATGCCATTTGCCACTACAGGCTCTCCATACCACTCCACGGAGAAAGAG	7333
7264	ACACGGATGCCATTTGCCACTACAGGCTCTCCATACCACTCCACGGAGAAAGAG	7323
7334	GACGGTGTCTGCAGAGAAATCCAAATGTCTTCTTCCCTTGGCGGAGTCTCGCACTAAGAC	7393
7324	GACGGTGTCTGCAGAGAAATCCAAATGTCTTCTTCCCTTGGCGGAGTCTCGCACTAAGAC	7383

Qy	7394	CTTCGGTAGCTCCGATCGTTCGGCCGTTTCATAGCGCACGGCGACCGCCCTTCCTGACCT	7453
Dd	7384	CTTCGGTAGCTCCGATCGTTCGGCCGTTTCATAGCGCACGGCGACCGCCCTTCCTGACGW	7443
Qy	7454	GGCTCCGACGACGGTGACAAAGGATCCGAGTTCGATCGTACCTCCATGCCGCCCT	7513
Dd	7444	GGCTCCGACGACGGTGACAAAGGATCCGAGTTCGATCGTACCTCCATGCCGCCCT	7503
Qy	7514	TGAAGGGAGCGGGGGACCCGATCTCAGCGACGGGCTTCGGTCTACCGTGAGTGAGGA	7573
Dd	7504	TGAAGGGAGCGGGGGACCCGATCTCAGCGACGGGCTTCGGTCTACCGTGAGTGAGGA	7563
Qy	7574	GGCTAGTGAGGATGTCGTCTGCTGCTCAATGTCCTATACGTGGACAGCGCCCTGATCAC	7633
Dd	7564	GGCTAGTGAGGACGTGCTGCTGCTCAATGTCCTATACGTGGACAGCGCTGATCAC	7623
Qy	7634	GCATGCGTGGGAGGAAAGTAAGCTGCCATCAACCCGTTGAGCAACTCTTTGCTGCG	7693
Dd	7624	GCATGCGTGGGAGGAAAGTAAGCTGCCATCAACCCGTTGAGCAACTCTTTGCTGCG	7683
Qy	7694	TCACCACACATGCTACGCCACACATCCCGACGCAAGCCTCCGGCAGAGAGAGGT	7753
Dd	7684	TCACCACACATGCTACGCCACACATCCCGACGCAAGCCTCCGGCAGAGAGAGGT	7743
Qy	7754	CACCTTTGACAGATTGCAAGTCTTCGATGATCATTTACCGGACAGTACTCAAGGAGATGAA	7813
Dd	7744	CACCTTTGACAGATTGCAAGTCTTCGATGATCATTTACCGGACAGTACTCAAGGAGATGAA	7803
Qy	7814	GGCGAAGCGTCCACAGTTAAGGCTTCTATCTATAGAGGAGGCTCCGAACTGAC	7873
Dd	7804	GGCGAAGCGTCCACAGTTAAGGCTTCTATCTATAGAGGAGGCTCCGAACTGAC	7863
Qy	7874	GGCCCCAATTCCGCCAATCCAAATTTGGCTATCGGGCAAGGACGTCCGGAACCTATC	7933
Dd	7864	GGCCCCAATTCCGCCAATCCAAATTTGGCTATCGGGCAAGGACGTCCGGAACCTATC	7923
Qy	7934	CAGCAGGCGGTTAACCATCCGCTCCGTTGGGAGGACTTGTCTGGAAGACACTGAAAC	7993
Dd	7924	CAGCAGGCGGTTAACCATCCGCTCCGTTGGGAGGACTTGTCTGGAAGACACTGAAAC	7983
Qy	7994	ACCAATTGACACCCATCATGGCAAAAGTAGGTTTTCTGCGTCCAAACAGAGAGGG	8053
Dd	7984	ACCAATTGACACCCATCATGGCAAAAGTAGGTTTTCTGCGTCCAAACAGAGAGGG	8043
Qy	8054	AGGCGCAAGCAGCTCGCTTATCGTATCCAGACCTGGAGTTCGTGTATGGAGAA	8113
Dd	8044	AGGCGCAAGCAGCTCGCTTATCGTATCCAGACCTGGAGTTCGTGTATGGAGAA	8103
Qy	8114	GATGGCCCTTTACGAGTGGTCTCCACCCCTTCCTCAGGCGGTGATGGGCTCCTCATACGG	8173
Dd	8104	GATGGCCCTTTACGAGTGGTCTCCACCCCTTCCTCAGGCGGTGATGGGCTCCTCATACGG	8163
Qy	8174	ATTTCAAATACTCCCAAGCAGCGGTTCGATTCCTGTGTAATACCTGGAATCAAAGAA	8233
Dd	8164	ATTTCAAATACTCCCAAGCAGCGGTTCGATTCCTGTGTAATACCTGGAATCAAAGAA	8223
Qy	8234	ATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGTACTCAACGGTCACTGAGAGTGA	8293
Dd	8224	ATGCCCTATGGGCTTCTCATATGACACCCGCTGTTTGTACTCAACGGTCACTGAGAGTGA	8283
Qy	8294	CAATTCGTTGAGGAGTCAATTTTACCATGTTGTGACTTGGCCCGGCGGCGCAGACAGGC	8353
Dd	8284	CAATTCGTTGAGGAGTCAATTTTACCATGTTGTGACTTGGCCCGGCGGCGCAGACAGGC	8343
Qy	8354	CATAAGGTCGCTCACAGACGGCTTTACATCCGGGGTCCCTGACTAACTCAAAAGGGCA	8413
Dd	8344	CATAAGGTCGCTCACAGACGGCTTTACATCCGGGGTCCCTGACTAACTCAAAAGGGCA	8403
Qy	8414	GAACTCGGTTATCGCCGGTGGCGCAAGTGGCGTGTGACGACTAGCTCGGTAATAC	8473
Dd	8404	GAACTCGGTTATCGCCGGTGGCGCAAGTGGCGTGTGACGACTAGCTCGGTAATAC	8463

Qy	8474	CCTCACATGTTACTTGAAGCCCACTGCAGCGCTGTCGAGCTGCAAAGCTCCAGGACTGCAC	8533
Dd	8464	CCTCACATGTTACTTGAAGCCCACTGCAGCGCTGTCGAGCTGCAAAGCTCCAGGACTGCAC	8523
Qy	8534	GATGCTCGTGAACGGGAGACGACCTTGTCTGTAATCTGTGAAGCGCGGAAACCCAGGAGGA	8593
Dd	8524	GATGCTCGTGAACGGGAGACGACCTTGTCTGTAATCTGTGAAGCGCGGAAACCCAGGAGGA	8583
Qy	8594	TGCGGCGGCTTACGAGCTTTCAGGAGGCTATGACTAGTATTTCCCGCCCCCCCCGGGGA	8653
Dd	8584	TGCGGCGGCTTACGAGCTTTCAGGAGGCTATGACTAGTATTTCCCGCCCCCCCCGGGGA	8643
Qy	8654	TCGCCCCCAACCAAGATPACGACCTGGAGCTGATTAACATCATGTTCCTCCAATGTGTCAGT	8713
Dd	8644	TCGCCCCCAACCAAGATPACGACCTGGAGCTGATTAACATCATGTTCCTCCAATGTGTCAGT	8703
Qy	8714	CGGCGACGATGTCATCTGGCAAAAGGTTACTACTCACCCTGACCGCCACCCACCCCT	8773
Dd	8704	CGGCGACGATGTCATCTGGCAAAAGGTTACTACTCACCCTGACCGCCACCCACCCCT	8763
Qy	8774	TGCACGGCTGCGTGGGAGACAGCTAGACACACTCCCAATCAACTCTTTGGCTTAGGCAATAT	8833
Dd	8764	TGCACGGCTGCGTGGGAGACAGCTAGACACACTCCCAATCAACTCTTTGGCTTAGGCAATAT	8823
Qy	8834	CATCATGTATGCGCCCACTTATGGGCAAGGATGATTTCTGATGACTCACTTTTCTCCAT	8893
Dd	8824	CATCATGTATGCGCCCACTTATGGGCAAGGATGATTTCTGATGACTCACTTTTCTCCAT	8883
Qy	8894	CTTTCTAGCTCAGAGCAACTTGAAGGCTTGGATTTGCAGATCTACGGGGCTTGCTA	8953
Dd	8884	CTTTCTAGCTCAGAGCAACTTGAAGGCTTGGATTTGCAGATCTACGGGGCTTGCTA	8943
Qy	8954	CTCCATTGAGCCACTTCACCTACCTCAGATCATTTGAACGACTCCATGCTTTAGCGGATT	9013
Dd	8944	CTCCATTGAGCCACTTCACCTACCTCAGATCATTTGAACGACTCCATGCTTTAGCGGATT	9003
Qy	9014	TACACTCCACAGTTACTCTCCAGGTGAGATCAATPAGGTTGGCTTCATGCTCAGGAACT	9073
Dd	9004	TACACTCCACAGTTACTCTCCAGGTGAGATCAATPAGGTTGGCTTCATGCTCAGGAACT	9063
Qy	9074	TGGGTTACCACTTGGCAACTCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACT	9133
Dd	9064	TGGGTTACCACTTGGCAACTCTGGAGACATCGGGCCAGAAAGTGTCCGCGCTAAGCTACT	9123
Qy	9134	GTCCCAGGGGGAGGCGCCACTTGTGGCAGATACCTCTTTAACTGGGCGAGTAAGGAC	9193
Dd	9124	GTCCCAGGGGGAGGCGCCACTTGTGGCAGATACCTCTTTAACTGGGCGAGTAAGGAC	9183
Qy	9194	CAAGCTTAAACTCACTCCAATCCCGGCGGTCCAGCTGGACTTGTCTGGCTGGTTCGT	9253
Dd	9184	CAAGCTTAAACTCACTCCAATCCCGGCGGTCCAGCTGGACTTGTCTGGCTGGTTCGT	9243
Qy	9254	CGCTGGTTACAGCGGGGAGACATATATACAGCGCTGTCTGTCGCGGACCCCGCTGGTT	9313
Dd	9244	CGCTGGTTACAGCGGGGAGACATATATACAGCGCTGTCTGTCGCGGACCCCGCTGGTT	9303
Qy	9314	TCCGTTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCAATTTACCTGCTCCCAACCGGATG	9373
Dd	9304	TCCGTTGTGCTACTCTCTACTTTCTGTAGGGGTAGGCAATTTACCTGCTCCCAACCGGATG	9363
Qy	9374	AACGGGAGCTTAAACACTCCAGGCTTTAAGCCATTTCTGTTTTTTTTTTTTTTTTTTTT	9433
Dd	9364	AACGGGAGCTTAAACACTCCAGGCTTTAAGCCATTTCTGTTTTTTTTTTTTTTTTTTTT	9423
Qy	9434	TT	9493
Dd	9424	TT	9483
Qy	9494	T 9494	
Dd	9484	T 9484	

```
RESULT 7
LOCUS HPCJ491 9460 bp RNA linear VRL 01-FEB-2000
DEFINITION Hepatitis C virus genome, complete sequence.
ACCESSION D10750
VERSION D10750.1 GI:221606
KEYWORDS C protein; E protein; E2 protein; NS1 protein; NS2 protein; NS3 protein; NS4 protein; NS5 protein.
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES: sgRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
1 (bases 1 to 9460)
Okamoto,H., Kojima,M., Okada,S., Yoshizawa,H., Iizuka,H., Tanaka,T., Muchmore,E.E., Peterson,D.A., Ito,Y. and Mishihiro,S.
Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability
Virology 190 (2), 894-899 (1992)
1325713
2 (bases 1 to 9460)
Okamoto,H.
Direct Submission
Submitted (18-MAR-1992) Hiroaki Okamoto, Jichi Medical School, Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi 329-04, Japan (E-mail:hokamoto@jichi.ac.jp, Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
Submitted (18-Mar-1992) to DDBJ by:
Hiroaki Okamoto
Immunology Division
Jichi Medical School
Kawachi-gun, Tochigi 329-04
Japan
Phone: 0285-44-2111 x3334
Fax: 0285-44-1557
Location/Qualifiers
1. .9460
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HC-J4"
/isolate="HC-J4/91"
/db_xref="taxon:11103"
342. .9374
/codon_start=1
/product="polyprotein precursor"
/protein_id="BAA01583.1"
/db_xref="GI:221607"
/translation="MSTNPKPQRTKNTNRPDQVDFPGGGQIVGVYLLPRGPRL
SVRAKTRTSRQRRQPIPKARQPEGRWAQPGYPWPLIGNEGWAGWVLEPRG
GPRMVGPTDPRRSNRLGKVIDTLTCGFADLMGYPLVGLAPLGAARALAHGVRVLED
GVNATCNLPGCSFIFLLALLSLTIPASAYEVNRVNSGIYHVNTDCNSIYVEAAD
VIHPTPCGVCVRGNSRCWVALTPTLAARDASVPTTIRRHVDLLVGTAAFCSAMV
VGLDGGISIFLVSLQFTSPRHETVQDCNSIYVGHVSGHRMADMMNNSPTALVV
SQLIRIPOAVDVVAGAHGWGLAGIAYISWGNWAKVLIIVALLPAGVDGATYSGGVA
GRITSGTSLFSSGASQIKLVNTNSWGHINRTALNCNDSLTFTGLAALFYTHKFNSS
GCPERMASCRPIDFGAGWGIITYTEPNSPDQPCYCHYAPRPGCIYPASQVCGPVYC
FTPSVAVGTIDRGVPTYSNGENETDMLNTRPPQGNWFGCTWMNSTGFTKTCG
PPCNIGGVNRTLCPTDCKPKHPEATVTKCGSGPWLTPCLVDYPVRLWHYPCPLNF
SIFKVRVYGVHRLNACNWTGERCNLEDRDRSELSPLLSTTEWQILPCAFTTL
PALSTGLIHQNIIVDQVLYGVGSFAFVSFAIKWEYIILLFLLADARVACNLMWLL
IAQEAALENLVLEMAASVAGAHGILSFLVFCAAWYIKGLAPGAAYAFYGVWPLL
LLALPPRAYALDEMAASCGGAVLVGLVFLTSPYKVFRLIIMWQYFTRAEAH
MQVWVPLNVKRGDAIILLTCAVHPELIIDITKGLAILGLPLVMOAGITRVPYFVR
AQQILRACMLRVAGGHYVOMAFKGLGALTGVYVNHLEPLRDWAHAGLRDLAVFE
PVFISMETKVIWGTADTAACGDIILGLIPVSARGKEIFLGPADLSLQGWRLAPIT
AYSQOTRGVLGCIITSUTGRDNKQVEGVQVSTATQSLATNGVCVNTYHVGAGSK
TLAGKPGITQMYNTDLDLVGWAQPGFARPCSCSSDLYLVLTTHADYIPVRRG
DRGSLSPRVSYLKGSGGGLLPCPSGVVGVFRAAVCTRGVAKAVDFIPVESMETT
MRSVFNTDNSPPAVPOTFOVHLHPTGSGSKTKVPAAYAAQYKVLVNLNPSVAATL
GFGAYMSKAHGDINIRGVRTIITGSSITYSTYKPLADGCGSGGADIIICECHS
TDSITLIGIGVLOQAGTAGARLVVLATATPFGSVTPVPHNIEIGISNNGEIPFGK
AIPTEAKIGRHLIFCHSKKCBELAKLGLGNVAYRGLDVSVPPIPGDVVVVA
TDALMTGTGDFDSVDCNTQVTQTFDPSLDPTFTIETTTPQDAVSRSQRGRGTGR
RSGIVRFVTPGERPSGMFDSVCEYDAGCAWVELTTPAETSVRLEAYLNPGLPVQC
```

```

DHLFEWSVFTCLTHIDAHFLSOTKQAGDNFYLVAQATVCARAQAAPPSPSDOMWKC
LIRLXPTLHGFTPLLIRLGAQVNEVILTHPIKTYIMACMSADLEVVTSTWLVGGVLA
ALAAVCLTGVSVIIVIRIILSGKPAVVPDREVLIOEFDEMEECASQLPIEBOGMLAE
QPKQALGLQATQKQAEAAAPVVEKRALETFWAKHMWNFISGQIYLAGUSTLPGN
PAIASLMAFTASITSLPTTQNTLLFNILGWNVAQAAPPASAFVAGIAGAAGVSI
GLRHVGDLAGYAGVAGALVAFKVMGVEVSTEDLVNLLPAILSPGALVGVVCAA
ILRRHVGPGEAGVOMNRLIAPASRGNHVSPTHYVPESDAAARVTQIILSSLTITOLK
RLHWINEDSCSTPCSGSLRWDMWICTVLTDTFKTLOSLLKPLRPVGPVFLSCQGYK
GVMRGDGMOTITPCGAQIAGHVKNGMRIVGPRTCSNTHWGTPIINAVTTCQCTPSP
APNYSRALRWAAEYEVTRVGDPHYVTGMTINDVKPCQCFAPFEFFTEVDGRLHR
YAPACKLLREDVTFQVGNLYKVSQCPCEPEPDVTVLTSMLTDPESHITAETAKRRL
ARGSPSLASSASQSLAPSKATCTTHHDSPADLIEANLLRWQEMGNIIRVESEN
KVWILDSPELHAEGDEREISVAAILRKSRKFPSPALPIWARPDVNPPLLESWMDPDY
VPPVHGCELPPTKAPPIPPERKKTVLTESNVSSALAEATKTFGSGSSGAVDSGT
ATALPDASDDGDKGSDVESYSMPLEGEPPDLSOGSMSTVEEASEDEVVCCSMS
DYHVRDLKEMKASTKRAKLSIEACKLTPHSAKSGFYGAKDVRNLSRAVNHIL
RSWEDLLEDTEPTDITIMAKSEVFCQPERKGRKPARLIIVFPDLGVVRCKMALYD
VVSTLPOAVMGSSYGYFOYSPKORVFLVNTWKSKKCPMGFSYDTRCFDSVTESDIRV
EESIYQCCDLAPEARQAIRSLTERLYIGPLTNSKONGCYRRCASGVLTTCGNTL
TCYLKATAACRAAKLQDCITMLVNGDDLVIIVICISAGTQEDAAALRAFTAMTRYASPPG
DPPOPEYDLELITSCSNVSVAHDSGRVYVLTDPITPLARAWETARHPINBWL
GNIMVAPTLWARMILMTFFSILLAQOELEKALDCQIYACYSIBPLDLPQIIEHLH
GLSAFTLHSYSGEINRVASCLKGLVPLRTWRHARSVRAKLSQGGRAATCGRYL
FNWAVTKLKLTFIPAAASQLDLSGWFVAGYSGDIYHSLSRARPRWFPLCLLLFVGV
GYLLPNR"
342. .914
/product="C protein"
915. .1490
/product="E protein"
1491. .2528
/product="NS1/E2 protein"
2529. .3359
/product="NS2 protein"
3360. .5186
/product="NS3 protein"
5187. .6380
/product="NS4 protein"
6381. .9371
/product="NS5 protein"
6005. .6104
/note="possible stem-loop structure in NS4 sequence"
9413. .9460
/note="T-stretch of 48 bp"

ORIGIN
Query Match 97.2%; Score 9327.4; DB 13; Length 9460;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 9376; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 1 GCCAGCCCCCTGATGGGGCGACATCCACCATGAATCACTCCCTGTGAGGAACACTAGT 60
Db 1 GCCAGCCCCCGATTGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAACACTAGT 60

Qy 61 TCTTACGCGAAGCGCTTAGCCATGGCGCTTAGTAGTGTGCTGACGCTCCAGGAC 120
Db 61 TCTTACGCGAAGCGCTTAGCCATGGCGCTTAGTAGTGTGCTGACGCTCCAGGAC 120

Qy 121 CCCCCCTCCGCGAGAGCCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180
Db 121 CCCCCCTCCGCGAGAGCCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180

Qy 181 GACGACCGGTCCTTTCTTGGATCAACCCGCTCAATCCCTGGAGATTTGGCGGTGCCCCC 240
Db 181 GACGACCGGTCCTTTCTTGGATCAACCCGCTCAATCCCTGGAGATTTGGCGGTGCCCCC 240

Qy 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGAAAGCCCTTGTGGTACTGCTCCGTATAGG 300
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGAAAGCCCTTGTGGTACTGCTCCGTATAGG 300

Qy 301 GTGCTTCGAGTCCCGGGAGGTCTCGTAGCCCGTGCACCATGACGACGAATCTCTAAC 360
Db 301 GTGCTTCGAGTCCCGGGAGGTCTCGTAGCCCGTGCACCATGACGACGAATCTCTAAC 360
```


Db 2521 |||||CAGACGCGCGTGTGCTGCTTGTGGATGATGCTGTGATAGCCAGGCTGAGGCCG 2580
Qy 2581 CTTTAGAGAACTTGGTGGTCTCAATGCGCGCTCCGTGGCCGAGCGCATGTTCTCT 2640
Db 2581 CTTTAGAGAACTTGGTGGTCTCAATGCGCGCTCCGTGGCCGAGCGCATGTTCTCT 2640
Qy 2641 CTTTCTTGTGTTCTTCTGCGCGCTCTGTACATTAAGGCGAGCTGGCTCCTGGGCGG 2700
Db 2641 CTTTCTTGTGTTCTTCTGCGCGCTCTGTACATTAAGGCGAGCTGGCTCCTGGGCGG 2700
Qy 2701 CGTATGCTTTTATGCGGTATGCGCGCTGCTGCTGCTCTCTACTGCGGTACACACAGAG 2760
Db 2701 CGTATGCTTTTATGCGGTATGCGCGCTGCTGCTGCTCTCTACTGCGGTACACACAGAG 2760
Qy 2761 CTTTACGCCCTTGGACCGGGAGATGGCTGCAATGCGTGGGGGTGCGGTTCTTGTAGGTCTGG 2820
Db 2761 CTTTACGCCCTTGGACCGGGAGATGGCTGCAATGCGTGGGGGTGCGGTTCTTGTAGGTCTGG 2820
Qy 2821 TATTCTTGACCTTGTCAACATCTACAAAGTGTTCCTACTAGGCTCATATGTTGTTAC 2880
Db 2821 TATTCTTGACCTTGTCAACATCTACAAAGTGTTCCTACTAGGCTCATATGTTGTTAC 2880
Qy 2881 AATACTTTATCACAGAGCGGCGCACATGCAAGTGTGGTCCCGCCCTCAACGTTT 2940
Db 2881 AATACTTTATCACAGAGCGGCGCACATGCAAGTGTGGTCCCGCCCTCAACGTTT 2940
Qy 2941 GGGAGGCGCGATGCCATCATCTCTCTACGTTGCGGTTTCATCAGAGTAAATTTTG 3000
Db 2941 GGGAGGCGCGATGCCATCATCTCTCTACGTTGCGGTTTCATCAGAGTAAATTTTG 3000
Qy 3001 ACATACCAAACCTCTGCTCGCATACTCGGCGGCTCATGTTGCTCCAGGCTGGCATAA 3060
Db 3001 ACATACCAAACCTCTGCTCGCATACTCGGCGGCTCATGTTGCTCCAGGCTGGCATAA 3060
Qy 3061 CGAGAGTCCGTTACTTGTGCGGCTCAAGGCTCATTCGTCGATGCTAGTACGAA 3120
Db 3061 CGAGAGTCCGTTACTTGTGCGGCTCAAGGCTCATTCGTCGATGCTAGTACGAA 3120
Qy 3121 AAGTCCGCGGGGTCAATTATGTCCAAATGTTCTTATGAAGCTGGGCGGCTGACAGGTA 3180
Db 3121 AAGTCCGCGGGGTCAATTATGTCCAAATGTTCTTATGAAGCTGGGCGGCTGACAGGTA 3180
Qy 3181 CGTACGTTTATAACCATCTTACCCCACTGCGGCACTGGGCCCAAGGCTCATACGAGCC 3240
Db 3181 CGTACGTTTATAACCATCTTACCCCACTGCGGCACTGGGCCCAAGGCTCATACGAGCC 3240
Qy 3241 TTGCGGTGGCGGTAGAGCCCGTCTCTTCTCCGCCATGGAGACCAAGGTCATACCTGGG 3300
Db 3241 TTGCGGTGGCGGTAGAGCCCGTCTCTTCTCCGCCATGGAGACCAAGGTCATACCTGGG 3300
Qy 3301 GAGCAGACCGCTGCTGTTGGGACATCATCTTGGGTCTACCGTCTCCGCCGAAGGG 3360
Db 3301 GAGCAGACCGCAGCATGTGGAGACATCATCTTGGGCTTACCGTCTCCGCCGAAGGG 3360
Qy 3361 GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGCAAGGTTGGGACTCCTTG 3420
Db 3361 GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGCAAGGTTGGGACTCCTTG 3420
Qy 3421 CGCCCATACCGCTACTCTCCAAACAAAGCGGGGCTACTTGGTTGATCATCATAGCC 3480
Db 3421 CGCCCATACCGCTACTCTCCAAACAAAGCGGGGCTACTTGGTTGATCATCATAGCC 3480
Qy 3481 TCACAGCGGGGACAGAACCAAGTCGAGGGGAGGTTCAAGTGGTTCTACCCCAACAC 3540
Db 3481 TCACAGCGGGGACAGAACCAAGTCGAGGGGAGGTTCAAGTGGTTCTACCCCAACAC 3540
Qy 3541 AATCTTCTCGGACCTGATCAACGCGGTGCTGGACTGTCTACCATGGCGTGGCT 3600
Db 3541 AATCTTCTCGGACCTGATCAACGCGGTGCTGGACTGTCTACCATGGCGTGGCT 3600
Qy 3601 CGAAGACCTTAGCCGGTCCAAAGGTCCAATCACCCAAATGTACCAATGTAGACCTGG 3660

Db 3601 CGAAGACCTTAGCCGGTCCAAAGGTCCAATCACCCAAATGTACCAATGTAGACCTGG 3660
Qy 3661 ACTCGTGGCTGGCAGGCGCCCCCGGGCGGCTCATGACACCATGAGCTGGCA 3720
Db 3661 ACTCGTGGCTGGCAGGCGCCCCCGGGCGGCTCATGACACCATGAGCTGGCA 3720
Qy 3721 GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCATTCGGTGGCGCGGAGGGG 3780
Db 3721 GCTCGGACCTTTACTTGGTCAAGACATGCTGATGTCATTCGGTGGCGCGGAGGGG 3780
Qy 3781 ACAGCAGGGGAAGTCTACTCTCCCCCAGCCCGCTCTCTACCTGAAGGGCTCTCGGGTG 3840
Db 3781 ACAGCAGGGGAAGTCTACTCTCCCCCAGCCCGCTCTCTACCTGAAGGGCTCTCGGGTG 3840
Qy 3841 GTCCATTGTTCCTTCGGGGCAGCTGTTGGGCTCTTCGGGCTGCTGTGTGACCC 3900
Db 3841 GTCCATTGTTCCTTCGGGGCAGCTGTTGGGCTCTTCGGGCTGCTGTGTGACCC 3900
Qy 3901 GGGGGTCCGGAAGGGGTGGACTTCATACCCGTTGAGTCTATGGAACCTACCATCGGT 3960
Db 3901 GGGGGTCCGGAAGGGGTGGACTTCATACCCGTTGAGTCTATGGAACCTACCATCGGT 3960
Qy 3961 CTCCGGTCTTTCAGACAACTCAACCCCCCGGCTGTACCGCAGACATTCCAAAGTGGCAC 4020
Db 3961 CTCCGGTCTTTCAGACAACTCAACCCCCCGGCTGTACCGCAGACATTCCAAAGTGGCAC 4020
Qy 4021 ATCTGACGCTCTACTGCGAGCGGCAAGACACCAAGTCCCGGCTGGTGTATGACGCC 4080
Db 4021 ATCTGACGCTCTACTGCGAGCGGCAAGACACCAAGTCCCGGCTGGTGTATGACGCC 4080
Qy 4081 AAGGTACAGGTGCTGCTCTGAAACCCGTCGTTGCCGACCTTAGGGTTTGGGGCGT 4140
Db 4081 AAGGTACAGGTGCTGCTCTGAAACCCGTCGTTGCCGACCTTAGGGTTTGGGGCGT 4140
Qy 4141 ATATGTCCAAAGGCACAGCTATCGACCTAACTCAGAACTGGGGTAAGAACCATTAACA 4200
Db 4141 ATATGTCCAAAGGCACAGCTATCGACCTAACTCAGAACTGGGGTAAGAACCATTAACA 4200
Qy 4201 CGGGCGGCTCCATTAGCTACTCCACTTGGCAAGTTCCTTGGCAGCGTGGTGTCTG 4260
Db 4201 CGGGCGGCTCCATTAGCTACTCCACTTGGCAAGTTCCTTGGCAGCGTGGTGTCTG 4260
Qy 4261 GGGCGGCTTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACTACCATCT 4320
Db 4261 GGGCGGCTTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACTACCATCT 4320
Qy 4321 TGGGCATCGGCACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGGCTGCTGTGCTG 4380
Db 4321 TGGGCATCGGCACAGTCTCTGACCAAGCGGAGACGGCTGGAGCGGCTGCTGTGCTG 4380
Qy 4381 CCACCGCTACACCTCGGGATCGGTTACCGTGGCACAACCCCAATATCGAGAAATAGGCC 4440
Db 4381 CCACCGCTACACCTCGGGATCGGTTACCGTGGCACAACCCCAATATCGAGAAATAGGCC 4440
Qy 4441 TGTCCAAATGAGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGGCCATCAAGG 4500
Db 4441 TGTCCAAATGAGAGATCCCTTCTATGGCAAGCCATCCCATTTGAGGCCATCAAGG 4500
Qy 4501 GGGGAGGCTATCTCATTTTCTGCCATTCGAAGAAATGTGACGAGCTCGCCGCAAGC 4560
Db 4501 GGGGAGGCTATCTCATTTTCTGCCATTCGAAGAAATGTGACGAGCTCGCCGCAAGC 4560
Qy 4561 TGACAGGCTCGGACTGAACGCTGTAGCATATTAACGGGGCTTGTGTGCTGTATAC 4620
Db 4561 TGACAGGCTCGGACTGAACGCTGTAGCATATTAACGGGGCTTGTGTGCTGTATAC 4620
Qy 4621 CGCCTATCGAGAGCTGTTGCTGGCAACAGACGCTCTAATGACGGGTTTACCCGGCG 4680
Db 4621 CGCCTATCGAGAGCTGTTGCTGGCAACAGACGCTCTAATGACGGGTTTACCCGGCG 4680
Qy 4681 ATTTTGACTCAGTGCATCGACTGCAATCATGTGTCCACGACAGTGCAGCTTACGTTGG 4740
Db 4681 ATTTTGACTCAGTGCATCGACTGCAATCATGTGTCCACGACAGTGCAGCTTACGTTGG 4740

QY 4741 ATCCACCTTACCAATTGAGACGACGACCGTGCCCAAGACGCGTGTGCGCTCGCAAC 4800
Db 4741 ATCCACCTTACCAATTGAGACGACGACCGTGCCCAAGACGCGTGTGCGCTCGCAAC 4800
QY 4801 GCGAGGTAGAACTGCGCAGGGTAGAGTGGCATCTACAGGTTGTGTGATCCAGAGAAAC 4860
Db 4801 GCGAGGTAGAACTGCGCAGGGTAGAGTGGCATCTACAGGTTGTGTGATCCAGAGAAAC 4860
QY 4861 GGCCTCGGGCATGTTGCATTTCTCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4920
Db 4861 GGCCTCGGGCATGTTGCATTTCTCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4920
QY 4921 GGTATGAGCTACGCGCTGAGACTCGGTTAGGTTGCGGCTTTACCTAAATACACAG 4980
Db 4921 GGTATGAGCTACGCGCTGAGACTCGGTTAGGTTGCGGCTTTACCTAAATACACAG 4980
QY 4981 GGTGCGGCTGTGCCAGGACCACTTGGAGTTCTGGGAGAGCGTCTTACAGGCTTCACCC 5040
Db 4981 GGTGCGGCTGTGCCAGGACCACTTGGAGTTCTGGGAGAGCGTCTTACAGGCTTCACCC 5040
QY 5041 ACATAGATGCCCATCTCTGTCTCCAGACTAAACAGCGCAGGAGAACCTTTTACCTGG 5100
Db 5041 ACATAGATGCCCATCTCTGTCTCCAGACTAAACAGCGCAGGAGAACCTTTTACCTGG 5100
QY 5101 TGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAA 5160
Db 5101 TGGCATATCAAGCTACAGTGTGCGCAGGGCTCAAGCTCCACCTCCATCGTGGGACCAA 5160
QY 5161 TGTGGAAGTGTCTCATACGGCTGAAACCTTACACTGACGCGGCAACACCCCTGTGTATA 5220
Db 5161 TGTGGAAGTGTCTCATACGGCTGAAACCTTACACTGACGCGGCAACACCCCTGTGTATA 5220
QY 5221 GGTAGGAGCCGTCCAAATAGAGTTCCTCCACACACCCCATAACTAAATACATCATGG 5280
Db 5221 GGTAGGAGCCGTCCAAATAGAGTTCCTCCACACACCCCATAACTAAATACATCATGG 5280
QY 5281 CATGATGTGCGCTGACCTGGAGTGTGCTACCTAGCACCTGGTGTGTGTGTGTGTGTGT 5340
Db 5281 CATGATGTGCGCTGACCTGGAGTGTGCTACCTAGCACCTGGTGTGTGTGTGTGTGTGT 5340
QY 5341 TTGCAAGTTTGGCGCATCTGCTGACGACAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5400
Db 5341 TTGCAAGTTTGGCGCATATGCTGTGACGACAGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 5400
QY 5401 TCTGTGCGGGAAGCCAGCTGTGTTCCCGACAGGAACTCTCTACAGGAGTTTGTGTGTGT 5460
Db 5401 TCTGTGCGGGAAGCCAGCTGTGTTCCCGACAGGAACTCTCTACAGGAGTTTGTGTGTGT 5460
QY 5461 AGATGGAAGTGTGCTTACAACTTCTTATCATCGAGCAGGAAATGCGAGTTCGCGGAGC 5520
Db 5461 AGATGGAAGTGTGCTTACAACTTCTTATCATCGAGCAGGAAATGCGAGTTCGCGGAGC 5520
QY 5521 AATTCAAGCAAGGCGCTCGGTTGTGCAAAAGGCGCAACGAGGAGGCTGTGTGTGTGTGTGT 5580
Db 5521 AATTCAAGCAAGGCGCTCGGTTGTGCAAAAGGCGCAACGAGGAGGCTGTGTGTGTGTGTGT 5580
QY 5581 CTCCCGTGTGAGTCCAAAGTGGCGAGCCCTTGTGAGACCTTGTGGGCGAAGCACAATGTGGA 5640
Db 5581 CTCCCGTGTGAGTCCAAAGTGGCGAGCCCTTGTGAGACCTTGTGGGCGAAGCACAATGTGGA 5640
QY 5641 ATTTATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTGTGCTGGAACCCCGGGA 5700
Db 5641 ATTTATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTGTGCTGGAACCCCGGGA 5700
QY 5701 TAGCATATTGATGCAATTTACAGCTTCTATCACTAGCCCGCTTACACCCCAAAACACCC 5760
Db 5701 TAGCATATTGATGCAATTTACAGCTTCTATCACTAGCCCGCTTACACCCCAAAACACCC 5760
QY 5761 TCTGTGTTAACTTTTGGGGGATGGGTGGCTGCCCAACTCGCTCTCCAGCGCTGCGT 5820
Db 5761 TCTGTGTTAACTTTTGGGGGATGGGTGGCTGCCCAACTCGCTCTCCAGCGCTGCGT 5820

QY 5821 CAGCTTTCTGTTGGCGCGGCATCGCGAGCGGCTGTGTGGCAGCATAGGCTTGGGAAG 5880
Db 5821 CAGCTTTCTGTTGGCGCGCGGCATCGCGAGCGGCTGTGTGGCAGCATAGGCTTGGGAAG 5880
QY 5881 TGTCTGTGGAATCTTGGCGGCTATGGGCGAGGGTAGCGGGCGCATCTGTGCGCTTTA 5940
Db 5881 TGTCTGTGGAATCTTGGCGGCTATGGGCGAGGGTAGCGGGCGCATCTGTGCGCTTTA 5940
QY 5941 AGGTATAGCGGCGAGGTGCCCTCCACCGAGGACCTGGTCAACTTACTCTCTGCGCATCC 6000
Db 5941 AGGTATAGCGGCGAGGTGCCCTCCACCGAGGACCTGGTCAACTTACTCTCTGCGCATCC 6000
QY 6001 TCTCTCTGTGCTGCTCGGCTCGTGTGCGCAGCAATACTTGTGCGTGGCAGCTGG 6060
Db 6001 TCTCTCTGTGCTGCTCGGCTCGTGTGCGCAGCAATACTTGTGCGTGGCAGCTGG 6060
QY 6061 GCCCGGAGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTGCTTCTGCGGGGTA 6120
Db 6061 GCCCGGAGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTGCTTCTGCGGGGTA 6120
QY 6121 ACCAGCTCTCCCTACGCACTATGTGCTGAGAGCGCTGCGAGCAGCTGCTCACTAGA 6180
Db 6121 ACCAGCTCTCCCTACGCACTATGTGCTGAGAGCGCTGCGAGCAGCTGCTCACTAGA 6180
QY 6181 TCCTCTTAGCTTTACCATCACTCAACTGTCTGAAGCGGCTCCACAGTGGATTAATGAG 6240
Db 6181 TCCTCTTAGCTTTACCATCACTCAACTGTCTGAAGCGGCTCCACAGTGGATTAATGAG 6240
QY 6241 ACTGCTTAGGCCATGTCTCGGCTCGTGGCTAAGGGATGTTTGGATTTGATGATGCG 6300
Db 6241 ACTGCTTAGGCCATGTCTCGGCTCGTGGCTAAGGGATGTTTGGATTTGATGATGCG 6300
QY 6301 TGTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCTGCGCGGTTTACCGGAGTCC 6360
Db 6301 TGTGACTGACTTCAAGACCTGGCTCCAGTCCAACTCTGCGCGGTTTACCGGAGTCC 6360
QY 6361 CTTTCTGTCTATGTCACCGCGGTACAAGGAGTCTGGCGGGGACGCGCATCATGCAAA 6420
Db 6361 CTTTCTGTCTATGTCACCGCGGTACAAGGAGTCTGGCGGGGACGCGCATCATGCAAA 6420
QY 6421 CCACCTGCCATGCGGAGCAGATGCGCGACATGTCAAAAACGGTTTCCATGAGGATCG 6480
Db 6421 CCACCTGCCATGCGGAGCAGATGCGCGACATGTCAAAAACGGTTTCCATGAGGATCG 6480
QY 6481 TAGGCTTAGAACCTGCGAGCAACGAGTGGCGAAGTTCCTCCCATCAACGATACACA 6540
Db 6481 TAGGCTTAGAACCTGCGAGCAACGAGTGGCGAAGTTCCTCCCATCAACGATACACA 6540
QY 6541 CGGACCTTGCACACCTCTCCCGCGCCCAACTATTCAGGCGCTATGCGGGTGGGTG 6600
Db 6541 CGGACCTTGCACACCTCTCCCGCGCCCAACTATTCAGGCGCTATGCGGGTGGGTG 6600
QY 6601 CTGAGGAGTACGTGGAGTTACGGTGTGGGATTTCCACTAGCTGACGGCATGACA 6660
Db 6601 CTGAGGAGTACGTGGAGTTACGGTGTGGGATTTCCACTAGCTGACGGCATGACA 6660
QY 6661 CTGACAACTGAAGTGGCCATGCGAGTTCGGCCCCCGAATTTCTACAGGAGTGGATG 6720
Db 6661 CTGACAACTGAAGTGGCCATGCGAGTTCGGCCCCCGAATTTCTACAGGAGTGGATG 6720
QY 6721 GAGTCCGCTTGCACAGGTACGCTCCGGCTGCAAACTCTTCTACGGGAGGACGTCACT 6780
Db 6721 GAGTCCGCTTGCACAGGTACGCTCCGGCTGCAAACTCTTCTACGGGAGGACGTCACT 6780
QY 6781 TCCAGGTCCGGCTCAACCAATATCTTGTGCGGTGCGAGTCCCATGCGAGCCGAAACCGG 6840
Db 6781 TCCAGGTCCGGCTCAACCAATATCTTGTGCGGTGCGAGTCCCATGCGAGCCGAAACCGG 6840
QY 6841 ACGTAAAGTGTCTTACTTCCATGCTCAGGATCCCTCCCACTTACAGGAGAGCGGCTA 6900
Db 6841 ACGTAAAGTGTCTTACTTCCATGCTCAGGATCCCTCCCACTTACAGGAGAGCGGCTA 6900
QY 6901 AGCGTAGGCTGGCTAGAGGGTCTTCCCGCTCTTTTAGCCAGCTCATCAGCTAGCTTGT 6960

Db	6901		AGCGTAGGCTGGCTAGAGGTCGCCCTCTTTAGCCAGCTCATCAGCTAGCCAGTTGT	6960
Qy	6961		CTGCGCTCTTTTGAAGGGACATGACACTACCCACCATGACTCCCGGACGCTGACTCA	7020
Db	6961		CTGCGCTCTTTTGAAGGGACATGACACTACCCACCATGACTCCCGGACGCTGACTCA	7020
Qy	7021		TCGAGGCCAACCTCTTTGTGGCGCAGAGATGCGGCGAAACATCATCTCGGTGGAGTCAG	7080
Db	7021		TCGAGGCCAACCTCTTTGTGGCGCAGAGATGCGGCGAAACATCATCTCGGTGGAGTCAG	7080
Qy	7081		AGAAATAGGTAGTAATCTTGGACTTTTCGAAACCGCTTTCACGCGGAGGGGATGAGAGGG	7140
Db	7081		AGAAATAGGTAGTAATCTTGGACTTTTCGAAACCGCTTTCACGCGGAGGGGATGAGAGGG	7140
Qy	7141		AGATATCCGTGCGGGGGAGATCCTGCGAAATCCAGGAAGTTCCCTCAGCGCTTGCCCA	7200
Db	7141		AGATATCCGTGCGGGGGAGATCCTGCGAAATCCAGGAAGTTCCCTCAGCGCTTGCCCA	7200
Qy	7201		TATGGGCACGCCCGGACTACAATCTCTCACTGTCTAGAGTCTCTGGAAGGACCCGAGCTACG	7260
Db	7201		TATGGGCACGCCCGGACTACAATCTCTCACTGTCTAGAGTCTCTGGAAGGACCCGAGCTACG	7260
Qy	7261		TCCTCTCGGTGGTACCGGATGCCATTCGCACTTACCAAGGCTCTCTCGAATACCACCTC	7320
Db	7261		TCCTCTCGGTGGTACCGGATGCCATTCGCACTTACCAAGGCTCTCTCGAATACCACCTC	7320
Qy	7321		CACGGAGAAAGAGCGGTTGTCTGACAGAACTCAATGTCTTCTGCTTGGCGGAGC	7380
Db	7321		CACGGAGAAAGAGCGGTTGTCTGACAGAACTCAATGTCTTCTGCTTGGCGGAGC	7380
Qy	7381		TCGCCCACTAAGACCTTCGGTAGCTCCGGATCGTGGCGCGTTGATPAGCGGCACGCGACCG	7440
Db	7381		TCGCCCACTAAGACCTTCGGTAGCTCCGGATCGTGGCGCGTTGATPAGCGGCACGCGACCG	7440
Qy	7441		CCCTTCTGACCTGGCTCCGACGACGCTGACAAAGGATCCGACGTTGAGTCGTACTCCT	7500
Db	7441		CCCTTCTGACCTGGCTCCGACGACGCTGACAAAGGATCCGACGTTGAGTCGTACTCCT	7500
Qy	7501		CCATGCCCTCTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGTCTTGGTCTA	7560
Db	7501		CCATGCCCTCTTGAAGGGAGCGGGGACCCCGATCTCAGCGACGGTCTTGGTCTA	7560
Qy	7561		CCGTGAGTGAGGAGCTAGTGAGGATGTCTGCTGCTCAATGTCTTATACGTGGACAG	7620
Db	7561		CCGTGAGTGAGGAGCTAGTGAGGACGTCTGCTGCTCAATGTCTTATACGTGGACAG	7620
Qy	7621		GCGCCCTGATCACGCCATGGCTCGGAGGAAAGTAAGCTGCCCCATCAACCCGTTGAGCA	7680
Db	7621		GCGCCCTGATCACGCCATGGCTCGGAGGAAAGTAAGCTGCCCCATCAACCCGTTGAGCA	7680
Qy	7681		ACTCTTTGCTGCTGCCACCAACATGTTCTACGCCCAACATCCCGGAGCGCAAGCCTCC	7740
Db	7681		ACTCTTTGCTGCTGCCACCAACATGTTCTACGCCCAACATCCCGGAGCGCAAGCCTCC	7740
Qy	7741		GGCAGAGAAGGTCACCTTTGACAGATGCAAGTCTCTGGATGATCATTTACCGGACGCTAC	7800
Db	7741		GGCAGAGAAGGTCACCTTTGACAGATGCAAGTCTCTGGATGATCATTTACCGGACGCTAC	7800
Qy	7801		TCAAGGAGATGAAGCGGAGCGTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGGAGG	7860
Db	7801		TCAAGGAGATGAAGCGGAGCGTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGGAGG	7860
Qy	7861		CCTGAGAGCTGACGCCCCCAATTCGGGCAAAATCCAAATTTGGCTATGGGCGAAAGNACG	7920
Db	7861		CCTGAGAGCTGACGCCCCCAATTCGGGCAAAATCCAAATTTGGCTATGGGCGAAAGNACG	7920
Qy	7921		TCGGGACCTATCAGCAGGCGGTTAACACATCCGCTCCGTGTTGGAGGACTTGTCTGG	7980
Db	7921		TCGGGACCTATCAGCAGGCGGTTAACACATCCGCTCCGTGTTGGAGGACTTGTCTGG	7980
Qy	7981		AAGACACTGAAACACCAATTGACACCCATCATGGCAAAAAGTGAGGTTTCTGCGTCC	8040

Db	7981		AAGACACTGAAACACCAATTGACACCCATCATGCGCAAAAAGTGAGGTTTCTGCGTCC	8040
Qy	8041		AACCAGAGAGGAGCGCCCAAGCCAGCTGCGCTTATCGTATTTCCAGACCTGGAGTTC	8100
Db	8041		AACCAGAGAGGAGCGCCCAAGCCAGCTGCGCTTATCGTATTTCCAGACCTGGAGTTC	8100
Qy	8101		GTGTATGCCAGAGATGGCCCTTTACGAGTGTCTCCACCTTCTCAGGCCGTGATGG	8160
Db	8101		GTGTATGCCAGAGATGGCCCTTTACGAGTGTCTCCACCTTCTCAGGCCGTGATGG	8160
Qy	8161		GCTCTCTCATACGGATTTCAATACTCCCCAAGCAGCGGGTFCGAGTTCTCTGGTGAATACCT	8220
Db	8161		GCTCTCTCATACGGATTTCAATACTCCCCAAGCAGCGGGTFCGAGTTCTCTGGTGAATACCT	8220
Qy	8221		GGAATCAAGAAATGCCCTTATGGGCTTCTCATATGACACCCGCTGTTTGTACTCAACGG	8280
Db	8221		GGAATCAAGAAATGCCCTTATGGGCTTCTCATATGACACCCGCTGTTTGTACTCAACGG	8280
Qy	8281		TCATCTGAGAGTGACATTCGTTGAGGAGTCAATTTTACCATGTTGTGACTTGGCCCCCG	8340
Db	8281		TCATCTGAGAGTGACATTCGTTGAGGAGTCAATTTTACCATGTTGTGACTTGGCCCCCG	8340
Qy	8341		AGCCAGACAGGCGCATAAGGTCCTCACAGAGCGGCTTTACATCGGGGGTCCCCGACTA	8400
Db	8341		AGCCAGACAGGCGCATAAGGTCCTCACAGAGCGGCTTTACATCGGGGGTCCCCGACTA	8400
Qy	8401		ACTCAAAAGGGCAGAACTGCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGACGACTA	8460
Db	8401		ACTCAAAAGGGCAGAACTGCGGTTATCGCGGTGCGCGCAAGTGGCGTGTGACGACTA	8460
Qy	8461		GCTGCGGTATACCTTCACATGTTTACTTGAAGCCACTGCGAGCCTGTCGAGCTGCAAGC	8520
Db	8461		GCTGCGGTATACCTTCACATGTTTACTTGAAGCCACTGCGAGCCTGTCGAGCTGCAAGC	8520
Qy	8521		TCAGAGACTGCAGATGCTCTGAAACGGAGACGACCTTGTCTGTTATCTGTGAAAGCGCG	8580
Db	8521		TCAGAGACTGCAGATGCTCTGAAACGGAGACGACCTTGTCTGTTATCTGTGAAAGCGCG	8580
Qy	8581		GAACCCAGGAGGATGCGCGGCCCTTACGAGCCTTTCAGCGAGGCTATGACTAGGTACTCCG	8640
Db	8581		GAACCCAGGAGGATGCGCGGCCCTTACGAGCCTTTCAGCGAGGCTATGACTAGGTACTCCG	8640
Qy	8641		CCCCCCCCGGGATCCGCCCCAACCAACGACCTGCGAGCTGATACATCATGTCTCT	8700
Db	8641		CCCCCCCCGGGATCCGCCCCAACCAACGACCTGCGAGCTGATACATCATGTCTCT	8700
Qy	8701		CCATGTGTCACTGCGGCAGATGCTATGCAAAAGGTTATCTACTCCTACCCGCTGACC	8760
Db	8701		CCATGTGTCACTGCGGCAGATGCTATGCAAAAGGTTATCTACTCCTACCCGCTGACC	8760
Qy	8761		CCACACCCCCCTTTGACGCGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAACTCTT	8820
Db	8761		CCACACCCCCCTTTGACGCGGCTGCGTGGGAGACAGCTAGACACACTCCAATCAACTCTT	8820
Qy	8821		GGCTAGGCAATATCATCATGTATCGGCCACCTTATGGGCAAGGATGTTCTGATGACTC	8880
Db	8821		GGCTAGGCAATATCATCATGTATCGGCCACCTTATGGGCAAGGATGTTCTGATGACTC	8880
Qy	8881		ACTTTTCTCCATCTCTTAGCTCAAGGCAACTTTGAAAAGCCCTGGATTGTGAGATCT	8940
Db	8881		ACTTTTCTCCATCTCTTAGCTCAAGGCAACTTTGAAAAGCCCTGGATTGTGAGATCT	8940
Qy	8941		ACGGGGCTTGCTACTCTCAATTGAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCATG	9000
Db	8941		ACGGGGCTTGCTACTCTCAATTGAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCATG	9000
Qy	9001		GTCTTAGCGCATTTACTTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060
Db	9001		GTCTTAGCGCATTTACTTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060
Qy	9061		GCCTCAGGAAACTTGGGGTACACCTTGGCAACTGAGACATCGGCCAGAGTGTC	9120
Db	9061		GCCTCAGGAAACTTGGGGTACACCTTGGCAACTGAGACATCGGCCAGAGTGTC	9120

9121 GGCCTAAGCTACTGTCCAGGGGGGAGGGCGGCACCTTGTGGCAGATACCTCTTTAACT 9180
|||||
9121 GGCCTAAGCTACTGTCCAGGGGGGAGGGCGGCACCTTGTGGCAGATACCTCTTTAACT 9180
|||||
9181 GGGCAGTAGGACCAAGCTTAACTCAGTCCCAATCCCGGGCGCGTCCAGCTGGACTGTT 9240
|||||
9181 GGGCAGTAGGACCAAGCTTAACTCAGTCCCAATCCCGGGCGCGTCCAGCTGGACTGTT 9240
|||||
9241 CTGGCTGGTTCTGCTCGCTGGTTACAGGGGGGAGACATATACACAGCTGTCTCGTGCC 9300
|||||
9241 CTGGCTGGTTCTGCTCGCTGGTTACAGGGGGGAGACATATACACAGCTGTCTCGTGCC 9300
|||||
9301 GACCCCGCTGGTTCCGTTGTCCCTACTCTACTTCTGTAGGGGTAGGCATTTACCTGCG 9360
|||||
9301 GACCCCGCTGGTTCCGTTGTCCCTACTCTACTTCTGTAGGGGTAGGCATTTACCTGCG 9360
|||||
9361 TCCCAACCGATGAACGGGGAGCTAACCACTCCAGGCTTTAAGCCATTTCTGTTTTTTT 9420
|||||
9361 TCCCAACCGATGAACGGGGAGCTAACCACTCCAGGCTTTAAGCCATTTCTGTTTTTTT 9420
|||||
9421 TTTTCTTCTTT 9457
|||||
9421 TTTTCTT 9457
|||||

RESULT 8
HPCJ483
LOCUS
DEFINITION
Hepatitis C virus genome, complete sequence.
ACCESSION
D13558 D01217
VERSION
D13558.1 GI:221604
KEYWORDS
C protein; E protein; E2 protein; NS1 protein; NS2 protein; NS3 protein; NS4 protein; NS5 protein.
SOURCE
Hepatitis C virus
ORGANISM
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE
1 (bases 1 to 9448)
Okamoto,H., Kojima,M., Okada,S., Yoshizawa,H., Iizuka,H., Tanaka,T., Muchmore,B.E., Peterson,D.A., Ito,Y. and Mishiro,S. Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability
Virology 190 (2), 894-899 (1992)
1325713
2 (bases 1 to 9448)
Okamoto,H.
Direct Submission
Submitted (17-Oct-1991) Hiroaki Okamoto, Jichi Medical School, Immunology Division; Minamikawachi-machi, Kawachi-gun, Tochigi 329-04 Japan (E-mail:hokamoto@jichi.ac.jp, Tel:0285-44-2111(ex.3334), Fax:0285-44-1557)
Submitted (17-Oct-1991) to DDBJ by:
Hiroaki Okamoto
Immunology Division
Jichi Medical School
Kawachi-gun
Tochigi 329-04
Japan
Phone: 0285-44-2111 x3334
Fax: 0285-44-1557.

FEATURES
source
1..9448
Location/Qualifiers
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="HC-J4"
/isolate="HC-J4/83"
/db_xref="taxon:11103"
342..9374
/codon_start=1
/product="polyprotein precursor"
/protein_id="BAA02756.1"
/db_xref="GI:221605"

CDS
342..914
/product="C protein"
915..1490
/product="E protein"
1491..2528
/product="NS1/E2 protein"
2529..3359
/product="NS2 protein"
3360..5186
/product="NS3 protein"
5187..6380
/product="NS4 protein"
6381..9371
/product="NS5 protein"
6005..6104
/note="possible stem-loop structure in NS4 sequence"
9413..9448
/note="T-stretch of 36 bp"

mat_peptide
mat_peptide
mat_peptide
mat_peptide
mat_peptide
mat_peptide
mat_peptide
stem_loop
misc_feature

ORIGIN

/translation="MSTNPKPQRKTKNTNRRPDQVXFPGGQIVGGVYLLPRGRPL
GVRATRTSRSPRGRQPIPKARPEGRAGQPGYPWPLYGNEGAGWAGLLSPRG
SRPSWPTDPRRSRNLGKVIDTLTCGFADLMGYIPLVAGPLGGAARALAGRVLED
MINHTPCVPCVEDNSRCHWALTPTLAARNSVPTTIRRHVDLLGAARAFCSAM
VGDLGCVFLVSQTLTFSRHEIVQDCNCSIEFGLSHGRMANDMMNNSTTALVV
SQLRIPQAVVMDVAGAHGVLAVLSVNWAKVLI VALLFAGVDGETTYSGGAA
SHTTSLIASLSPGASQRIQLVNTNGSHNRTALNCNDSLHTFGALFTHFRNSS
CTPRMASCRIPIWFAQGWGITVTEPDSQDRQWHPAPRCGIPVASQVSTGTCG
FTSPVVGTTDRSGVPTYSWGENETDMLNNTRPQGNWFGCTWMNSTGFTKTCG
PPCNIGVGNTLTCPTDPRKHPEATYTKCGSGPMLTPRCLVDPYVRLWHPCTEF
SIFKRVYGVGVEHRLNAACNWTGERCNLDRDRSELSPLLTSTTEHQIILCAFTTL
PALSTGLIHLQNIQVQVLYGVGSFAVSFAIKWEYIILLFLLADARVACLAQMLLL
IAQAALAEHLVNLNAASVAGAHGILSLVFFCAAWYIKGRFAAYLAQFYGVMLLL
LIALPRAVAMDREMAASCGVALVGLVLTSLPYKVFTRIIMLQYFTRAEALH
MOQVPLNVRGGGRDAIILLTCAVHPELIDITKLLAILGLMVAQVDFIIRVPYFVR
AQGLRACMLVRKVGAGHYQMAFMPKLGALTGYVYVNHILPLRMAHAGLRDLAVFE
PVFSDMETITWGAADTACGDIILGLVPSARRGKEIFLGPADLSDEQGRWLLAPIT
AYSQTRGVLCGIITSLTRDKNOGEVQVSTATOSFLATCINGVCVTVYHAGSK
TLAGPKGPIQMYTNVDLDLVQMAPPGARSMTPTCTGSSDLYLTVRHADYIVRRRG
DSRGLSPRPVSYLKGSSGGLLCPGSHVGVFRAAVCTRGAKAVDFIIPVESMETT
MRSFVFTDSSPPAVPOTFOVAHLHAPTGSGKSTKVPAAVAAQGVKVLVLPVAAITL
GFGAYMPKAGIDENIRGTITGTGPIYTYGKFLADGCGSGGAYDIIICBECHS
TDSITLIGITVLDAQETAGARLVVLATATPPGVTVPHPNIEELGNSNGEIPYK
AIPTEAIKGGRHILIFCHSKKCDLAAKLTGLNLNAVAYRGLDVSVPITPGDVAVVA
RSGIYRVFTGERSGMDSSVLCYDAGCAWELTPTAETSRLRAYLNTPLPVCQ
TDALMTGTGDFDSVIDCNTCTQTVDFSLDPTTETITTPQDAVSRORRGTRG
DHLFESVETGLSHIDAHFLSOTKOAGDNFPYLVAQATVCARAQAPPSDWDQWKC
LIRUKPTHTGPTPLLYRLGAVQNEVTLTHPTIKYIMACMSADLEVTSTVILGVGLA
ALAAQYSGTVIVGRIILSGKPAVDPDREVLVOEFDEMEBECASQLPIISQGMOLAE
QFKQALGLQTATKQAEAAAPVVESSKRALETFWAKHMNFISGIQYLAGLSTLPGN
PAIASMALFTASITSPITNTQNTLLFNILGVMVAQALAPPASASAFVGAAGAAVGS
GLGKVLVDILAGVAGVALVAFKVMSEVPSTEDLVNLLPAILSPALVGVVCAA
ILRHVPGEGAVQVMNRLIAFASGNHVSPTHVPESDAAARVTOILSLTITOLLK
RLHWNEDGCTPCSGSLRWDWDCITVLDFTKMLQSKLLPLRLPVFPLSCQGYK
GVMRGDGMOTPCGAQIAGHVRKSNRIVGPRCTCNTWHGTGTFINATYTGCPCTSP
APNTSRALMRVAEEYVEVTRGDFHYVTGMTDNTKVCQCPAPPEFTVEDVRLHR
YAPACKPLLRDVAFLGVLNOYLVGSQLPCPEPDVTLTSMLTDPESHITAEAKRL
ARGSPSLASSASQLSAPLKATCTTHDSPDADLIENLLRMEMPENITRVSEEN
KVILDSFEPHAGSDEREISVAAEILKRKSPSALPIWARPDYNPLLSKWDPDY
VPPVHCGPLPTKAPPIPPRRKTRVLTESNVSSALAEATKATFFSSGGSADVSGT
ATAPLQASDDGDKGSVSYSPMPLEGEFDPLSDGSWSTVSEESSEDVCCSMS
YTWGALITPCAASEKLPINPLSNLRLHNMVYATTSRSASLRQKVTFDRLQVLD
DHYRDLKEMAKASTVKALLSIEACKLTPHSAKSGFGYKDKVRLNLSRAVNH
RSVWEDLLEDTEPIDTTIMAKNEVCQPEKGRKPARLIVFDPLGVRCVKMALYD
VSTLPQAVMSGSYFQYSPKQVVELNWTWKKCPMGFSYDTFCFSTVETSDIRV
EESIYQCCDLAPRAQIRSLTERLIYIGPLTNSKGQNGRRCRASGVLTISCCNTL
TCYLKATAACRAALKQDCTMLVNGDDLVI CESAGTQEDAAALRVFTSAMTRYAPP
DPQPEYDLELITSCSNVSVAHDAKGRVYLTTRDPTIPLARAWEATHTPPVNSWL
GNITWYAPALWARMILMTHFSLIAQLQALDQCIYGCYSTEPDLDPQIIBRLH
GLSAFSLHSYSPGEINRVASCLRKLGVPLRVWRHARSVRKLLSQGGRATCGRYL
FNWAVRTKLKLTPIPAASQLDLSGFWFVAGYSGGDIYHLSLRARPRWFLLCULLLVGVG
GIYLLPNR"

Query Match		95.8%;	Score 9190.4;	DB 13;	Length 9448;
Best Local Similarity		98.3%;	Pred. No. 0;		
Matches 9287;		Conservative	0;	Mismatches 161;	Indels 0; Gaps 0;
QY	1	GCCAGCCCCCTGATGGGGCGGACACTCCACCATGAATCACTCCCTGTGTGAGGAATCTACTG	60		
DB	1	GCCAGCCCCCGATTGGGGGCGGACACTCCACCATAGATCACTCCCTGTGTGAGGAATCTACTG	60		
QY	61	TCCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTGAGCGCTTCCAGGAC	120		
DB	61	TCCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTGAGCGCTTCCAGGAC	120		
QY	121	CCCCCTCCCGGAGAGCCATAGTGTCTGTGCGAAACCGGTGAGTACACCGGAATTCGCCAG	180		
DB	121	CCCCCTCCCGGAGAGCCATAGTGTCTGTGCGAAACCGGTGAGTACACCGGAATTCGCCAG	180		
QY	181	GACGACCGGGTCTCTTTCTTGGATCAACCCGCTCAATGCTTGGAGATTTGGCGGTGCCCCC	240		
DB	181	GACGACCGGGTCTCTTTCTTGGATCAACCCGCTCAATGCTTGGAGATTTGGCGGTGCCCCC	240		
QY	241	GCAGAGCTCTAGCCGAGTAGTGTGGGTTCGGAAGGGCTTGTGTACTGTCCCTGATAGG	300		
DB	241	GCAGAGCTCTAGCCGAGTAGTGTGGGTTCGGAAGGGCTTGTGTACTGTCCCTGATAGG	300		
QY	301	GTGCTTCGAGTGCCCCGGGAGGCTCTGTAGACCGGTGCACCATGAGCAGCAATCTCTAAAC	360		
DB	301	GTGCTTCGAGTGCCCCGGGAGGCTCTGTAGACCGGTGCACCATGAGCAGCAATCTCTAAAC	360		
QY	361	CTCAAGAAAACCAAACTTAACACCAACCGCCGCCACAGAGACGTCAAGTTCCCGGGCG	420		
DB	361	CTCAAGAAAACCAAACTTAACACCAACCGCCGCCACAGAGACGTCAAGTTCCCGGGCG	420		
QY	421	GTGGTCAGATCGTTGTGTGAGTTTACTGTGTCCGCGCAGGGGCCCCAGGTTGGGTGTGC	480		
DB	421	GTGGTCAGATCGTTGTGTGAGTTTACTGTGTCCGCGCAGGGGCCCCAGGTTGGGTGTGC	480		
QY	481	GGCGACTAGGAAGGCTTCCGAGCGGTGCGAACTCGTGTGAAGGCGACACCTATCCCAA	540		
DB	481	GGCGACTAGGAAGACTTCCGAGCGGTGCGAACTCGTGTGAAGGCGACACCTATCCCCA	540		
QY	541	AGGCTCCGACCCGAGGGCAGGGCTTGGGCTCAGCCCGGGTACCCCTTGGCCCTCTATG	600		
DB	541	AGGCTCCGCGACCCGAGGGCAGGGCTTGGGCTCAGCCCGGGTACCCCTTGGCCCTCTATG	600		
QY	601	GCAATGAGGGCTTGGGTGGGCAAGGATGGCTCTGTCAACCGCGGCTCCCGGCTAGTT	660		
DB	601	GCAATGAGGGCTTGGGTGGGCAAGGATGGCTCTGTCAACCGCGGCTCCCGGCTAGTT	660		
QY	661	GGGGCCCCACGACCCCGCGGTAGGTGCGGTAACTTGGGTAAAGTCAATCGATACCCCTTA	720		
DB	661	GGGGCCCCACGACCCCGCGGTAGGTGCGGTAACTTGGGTAAAGTCAATCGATACCCCTTA	720		
QY	721	CATCGGGTTCGCGGATCTCATGGGTACAACTTCGCTGTCGCGCCCCCTTAGGGGGCG	780		
DB	721	CATCGGGCTTCGCGGATCTCATGGGTATATTCGCTGTCGCGCCCCCTTAGGGGGCG	780		
QY	781	CTGCCAGGGCTTGGCACACGGTGTCCGGGTTCGTGAGGACGGCGTGAATATGCAACAG	840		
DB	781	CTGCCAGGGCTTGGCACACGGTGTCCGGGTTCGTGAGGACGGCGTGAATATGCAACAG	840		
QY	841	GGAACTTCCCGGTGTCTTTCTCTATCTTCTCTTTGGCTCTGCTGTCTGTGTGACCA	900		
DB	841	GGAACTTCCCGGTGTCTTTCTCTATCTTCTCTTTGGCTCTGCTGTCTGTGTGTGACCA	900		
QY	901	TCCAGCTTCGGCTTATGAATGCGCAACGTGTCCGGATATACCATGTACGAAACGACT	960		
DB	901	TCCAGCTTCGGCTTATGAATGCGCAACGTGTCCGGATATACCATGTACGAAACGACT	960		
QY	961	GCTCCAACTCAAGCATTTGTATGAGGACGGGACGTGATCATGCATACTCCCGGGTGGC	1020		
DB	961	GCTCCAACTCAAGCATTTGTATGAGGACGGGACATGATCATGCATACTCCCGGGTGGC	1020		
QY	1021	TGCCCTGTGTTTCAGGAGGGTAAACAGCTCCCGTGTGTGGGTAGCGCTCACTCCCAACGCTC	1080		

DB	1021	TGCCCTGTGTTTCGGGAGGACAACAGCTCCGTTGCTGGGTAGGCTCACTCCACGCTCG	1080
QY	1081	CGGCCAGGAATGCAAGCGTCCCACCTACGACAATACGACGCCACGTCTGCTCGTTG	1140
DB	1081	CGGCCAGGAATGCAAGCGTCCCACCTACGACAATACGACGCCACGTCTGCTCGTTG	1140
QY	1141	GGACGGCTGTTTCTGCTCGGTATGTACGTGGGGATCTCTCGGATCTATTTCCTCG	1200
DB	1141	GGGCGCTGCTTCTGCTCGGTATGTACGTGGGGATCTCTCGGATCTATTTCCTCG	1200
QY	1201	TCCTCCAGCTGTTTCACTCTTCGCTCGCGCATGAGACAGTGCAGGACTGCAACTGCT	1260
DB	1201	TCCTCCAGCTGTTTCACTCTTCGCTCGCGCATGAGACAGTGCAGGACTGCAACTGCT	1260
QY	1261	CAATCTATCCCGGCCATGTATCAGGTCAACGATGGCTTGGGATATGATGATGACTGGT	1320
DB	1261	CAATCTATCCCGGCCATGTATCAGGTCAACGATGGCTTGGGATATGATGATGACTGGT	1320
QY	1321	CACCTACACAGCCCTAGTGTGCGAGTTGCTCCGGATCCCAAGCTGTCTGTGACA	1380
DB	1321	CACCTACACAGCCCTAGTGTGCGAGTTGCTCCGGATCCCAAGCTGTCTGTGACA	1380
QY	1381	TGTTGGCGGGGCCACTGGGGAGTCTTGGCGGCTTGGCTACTATTCCATGTAGGGA	1440
DB	1381	TGTTGGCGGGGCCACTGGGGAGTCTTGGCGGCTTGGCTACTATTCCATGTAGGGA	1440
QY	1441	ACTGGGCTAAGGTTCTGATTTGTCGCTACTCTTTCGCGGCTTACGCGGAGACCCACA	1500
DB	1441	ACTGGGCTAAGGTTCTGATTTGTCGCTACTCTTTCGCGGCTTACGCGGAGACCTACA	1500
QY	1501	CGACGGGAGGTTGGCGGCCACACACCTTCGGGTTCAGTCCCTTTCTCATCTCGGG	1560
DB	1501	CGTGGGGGGGGCGGCGAGCCACACACCTTCAGCTTCGCTCCCTTTCTTCTCCTCGGG	1560
QY	1561	CGTCTCAGAAATCCAGCTTGTGAATACCAACGCGACTGCGCATCAACAGGACTGCC	1620
DB	1561	CGTCTCAGAGATCCAGCTTGTGAATACCAACGCGACTGCGCATCAACAGGACTGCC	1620
QY	1621	TAAATTCGAATGACTCCCTCCAACTGGGTTCTTTCGCGGCTGTTTTAGCGACACAAGT	1680
DB	1621	TAAATTCGAATGACTCCCTCCACACTGGGTTCTTTCGCGGCTGTTTTACACACACAAGT	1680
QY	1681	TCAACTCGTCCGGTCCCGGAGGCGATGGCGAGCTCCCGGCCCATTCAGTGTTCGCC	1740
DB	1681	TCAACTCGTCCGGTCCCGGAGGCGATGGCGAGCTCCCGGCCCATTCAGTGTTCGCC	1740
QY	1741	AGGGTGGGGCCCCCATCACCTATACCTAAAGCTAACAGCTCGGATCAGAGGCTTTATGCT	1800
DB	1741	AGGGATGGGGCCCCCATCACCTATACCTAAAGCTAACAGCTCGGATCAGAGGCTTTATGCT	1800
QY	1801	GGCAATACGGGCTTCGACCGGTGTGTTGCTTACCGGCTCGCAGGTGTGGTCCAGTGT	1860
DB	1801	GGCAATACGGGCTTCGACCGGTGTGTTGCTTACCGGCTCGCAGGTGTGGTCCAGTGT	1860
QY	1861	ATTGTTTCAACCCCAAGCCCTGTTGTGTGGGACCAACGATTCGTTCCGTTGCTCTAGT	1920
DB	1861	ATTGTTTCAACCCCAAGCCCTGTTGTGTGGGACCAACGATTCGTTCCGTTGCTCTAGT	1920
QY	1921	ATAGCTGGGGGGAATGAGACAGCTGATGCTTCAACACACGCGTCCGCCACAAG	1980
DB	1921	ATAGCTGGGGGGAATGAGACAGCTGATGCTTCAACACACGCGTCCGCCACAAG	1980
QY	1981	GCAACTGTTTCGGCTGTACATGATGATAGTACTGGGTTCACTAAGAAGTTCGGAGGTC	2040
DB	1981	GCAACTGTTTCGGCTGTACATGATGATAGTACTGGGTTCACTAAGAAGTTCGGAGGTC	2040
QY	2041	CCCCGTGTAAACATCCGGGGGGTCCGTAACCGCACTTTGATGTGCCCCACGAGCTGCTTC	2100
DB	2041	CCCCGTGTAAACATCCGGGGGGTCCGTAACCGCACTTTGATGTGCCCCACGAGCTGCTTC	2100
QY	2101	GGAAAGCAACCCCGAGGCTACTTACACAAATGTGCTCGGGGCTTGGTTGACACTAGT	2160

Db 2101 GGAAGCA CCCCAGGCTACTTACAAAAATGTGGCTCGGGGCCCTGGTTGACACTAGGT 2160
Qy GCCTAGTAGACTACCATACAGGCTTTGGCACTACCCCTGCACCTCTCAATTTTTCATCT 2220
Db GCCTAGTAGACTACCATACAGGCTTTGGCACTACCCCTGCACCTTTCAATTTTTCATCT 2220
Qy TTAAGGTTAGGATGATGTGGGGCGTGGAGCAGAGCTCAATGCCGATGCAATTGGA 2280
Db TTAAGGTTAGGATGATGTGGGGCGTGGAGCAGAGCTCAATGCCGATGCAATTGGA 2280
Qy CTCGAGGAGAGCGCTGTAACCTTGGAGGACAGGATAGGTCAAGCTCAGCCGCTGTCTGC 2340
Db CTCGAGGAGAGCGCTGTAACCTTGGAGGACAGGATAGGTCAAGCTCAGCCGCTGTCTGC 2340
Qy TGTCTACAAACAGAGTGGAGTACTGCCCTGTGCTTTTACCAACCTACCGGCTTTATCCA 2400
Db TGTCTACAAACAGAGTGGAGTACTGCCCTGTGCTTTTACCAACCTACCGGCTTTTGTCCA 2400
Qy CTGGTTTGTATCCATCTCCATCAGAACATCGTGGAGCTGCAATACCTGTACGGTGTAGGGT 2460
Db CTGGTTTGTATCCATCTCCATCAGAACATCGTGGAGCTGCAATACCTGTACGGTGTAGGGT 2460
Qy CAGCGTTTGTCTCTCTTTCGAATCAAAATGGAGTACATCCTGTGTCTTTTCTCTCTCTGG 2520
Db CAGCGTTTGTCTCTCTTTCGAATCAAAATGGAGTACATCCTGTGTCTTTTCTCTCTCTGG 2520
Qy CAGACGCGCGCTGTGCTGCTTGTGGATGATGCTGTGATAGCTAGCCAGGCTGAGGCGG 2580
Db CAGACGCGCGCTGTGCTGCTTGTGGATGATGCTGTGATAGCTAGCCAGGCTGAGGCGG 2580
Qy CTTTATAGAACTTGTGTGCTCTCAATGCGGCTCGTGGCCGAGCGCATGATTTCTCT 2640
Db CTTTATAGAACTTGTGTGCTCTCAATGCGGCTCGTGGCCGAGCGCATGATTTCTCT 2640
Qy CTTTCTTGTGTCTCTTCTGCGCGCTGTGTA CATTAAGGGCAGGCTGTGCTCTGGGGCGG 2700
Db CTTTCTTGTGTCTCTTCTGCGCGCTGTGTA CATTAAGGGCAGGCTGTGCTCTGGGGCGG 2700
Qy CGTATGCTTTTATGCGGTATGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Db CGTATGCTTTTATGCGGTATGCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
Qy CTTTACGCCCTTGGACCGGGAGATGGTGTGCACTCGTGGGGGGTGGTCTTCTGTAGTCTCTGG 2820
Db CTTTACGCCATGACCGGGAGATGGTGTGCACTCGTGGGGGGTGGTCTTCTGTAGTCTCTGG 2820
Qy TATCTCTGACCTTGTACCATACTACAAAGTGTTCCTACTAGGCTCATATGGTGGTTAC 2880
Db TATCTCTGACCTTGTACCATACTACAAAGTGTTCCTACTAGGCTCATATGGTGGTTAC 2880
Qy AATACTTTATACCCAGAGCGGAGGCGCACATGCAAGTGGGTCCCCCCTCAACGTTTC 2940
Db AATACTTTATACCCAGAGCGGAGGCGCACATGCAAGTGGGTCCCCCCTCAACGTTTC 2940
Qy GGGAGGCGCGATGCCATCATCTCTCACTGTGCGGTTCATCCAGAGTTAAATTTTGG 3000
Db GGGAGGCGCGATGCCATCATCTCTCACTGTGCGGTTCATCCAGAGCTAAATTTTGG 3000
Qy ACATCAACAAATCTCTGTCTGCGCATACTCGGCCGCTCATGTGTCTCAGGCTGGCATAA 3060
Db ACATCAACAAATCTCTGTCTGCGCATACTCGGCCGCTCATGTGTCTCAGGCTGGCATAA 3060
Qy CGAGAGTCCGTACTCTGTGCGGCTCAAGGGCTCATTCGTGCATGCAATGTAGTGCAGAA 3120
Db CGAGAGTCCGTACTCTGTGCGGCTCAAGGGCTCATTCGTGCATGCAATGTAGTGCAGAA 3120
Qy AAGTCCCGGGGTCAATATGTCCAAATGGTCTTCAATGAAGCTGGGCGGCTGACAGGTA 3180
Db AAGTCCCGGGGTCAATATGTCCAAATGGTCTTCAATGAAGCTGGGCGGCTGACAGGTA 3180
Qy CGTACGTTTATAACCATCTTACCCCACTGCGGACTGGGCCCAACGCGGCTCTACGAGCC 3240
Db CGTACGTTTATAACCATCTTACCCCACTGCGGACTGGGCCCAACGCGGCTCTACGAGCC 3240

Qy 3241 TTGCGTGGGGTAGAGCCCGTCTGTCTTCTCCGCCATGGAGCAAGGTCTATCACTGGG 3300
Db TTGCGTGGCAGTAGAGCCCGTCTGTCTTCTCCGACATGGAGCAAGGTCTATCACTGGG 3300
Qy GAGCAGACACCGCTCGTGTGGGACATCATCTTTGGGTCTACCCGTCTCCGCCGGAAGG 3360
Db GAGCAGACACCGCAGCGTGTGGGACATCATCTTTGGGTCTACCCGTCTCCGCCGGAAGG 3360
Qy GGAAGGAGATATTTTGGGACCGGCTGATAGTCTCGAAGGCAAGGTTGGGACTTCCTTG 3420
Db GGAAGGAGATATTTTGGGACCGGCTGATAGTCTTTGAAGGCAAGGTTGGGACTTCCTTG 3420
Qy CGCCCATCA CGGCCCTACTCCCAACAAACCGGGGCGTACTTTGGTTGATCATCACTAGCC 3480
Db CGCCCATCA CGGCCCTACTCCCAACAAACCGGGGCGTACTTTGGTTGATCATCACTAGCC 3480
Qy TCACAGCGCGGACAAAGAACAGAGTCTGAAAGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
Db TCACAGCGCGGACAAAGAACAGAGTCTGAGGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
Qy AATCTTCTCTGGCGACCTGCAATCAACCGCGTGTCTGACTGTCTACCATGCGCTGGCT 3600
Db AATCTTCTCTGGCGACCTGCAATCAACCGCGTGTCTGACTGTCTACCATGCGCTGGCT 3600
Qy CGAAGACCTTAGCCGGTCCAAAAGGTCCAAATCACCACAAATGTACCAATGTAGACCTGG 3660
Db CGAAGACCTTAGCCGGTCCAAAAGGTCCAAATCACCACAAATGTACCAATGTAGACCTGG 3660
Qy ACTCTGTCTGGTGTGCGAGCGGCCCGCGGCGCTCTCATGACACCATGTCAGACTGTGCA 3720
Db ACTCTGTCTGGTGTGCGAGCGGCCCGCGGCGCTCTCATGACACCATGTCAGACTGTGCA 3720
Qy GCTCGGACCTTTACTTGGTACAGAGACATGCTGATGTCAATTCGGTGTGCGCGGAGCGG 3780
Db GCTCGGACCTTTACTTGGTACAGAGACATGCTGATGTCAATTCGGTGTGCGCGGAGCGG 3780
Qy ACAGCAGGGGAAGTCTACTCTCCCGCCAGGCGCGTCTCTACTCTGAAAAGGCTCTCTCGG 3840
Db ACAGCAGGGGAAGTCTACTCTCCCGCCAGGCGCGTCTCTACTCTGAAAAGGCTCTCTCGG 3840
Qy GTCCATGCTTTTGGCTTTCGGGGCACGTTGTGGGCGTCTTCCGGGCTGCTGTGTGCAACC 3900
Db GTCCATGCTTTTGGCTTTCGGGGCACGTTGTGGGCGTCTTCCGGGCTGCTGTGTGCAACC 3900
Qy GGGGGTTCGGAAGCGGTGCACTTATACCGGTGTAGTCTTGAAGTACCATGCGGT 3960
Db GGGGGTTCGGAAGCGGTGCACTTATACCGGTGTAGTCTTGAAGTACCATGCGGT 3960
Qy CTCGGTCTTTCACAGACAACTCAACCCCCCGGCTGTACCGCAGACATTTCCAAGTGGCAC 4020
Db CTCGGTCTTTCACAGACAACTCATCCCCCGGCTGTACCGCAGACATTTCCAAGTGGCAC 4020
Qy ATCTGACGCTCTTACTGTGCGAGCGGCAAGAGAACAAAGTGGCGCTGGGTATGAGGCC 4080
Db ATCTGACGCTCTTACTGTGCGAGCGGCAAGAGCACTAAAGTGGCGCTGGGTATGAGGCC 4080
Qy AAGGTACAAAGTGTCTGCTGTAACCGGTGGTGGCGCCACCTTAGGTTTGGGGCGT 4140
Db AAGGTACAAAGTGTCTGCTGTAACCGGTGGTGGCGCCACCTTAGGTTTGGGGCGT 4140
Qy ATATGTCAAGGCAACAGCTTACACCTTAACATCAGAACTGGGGTAAGGACCATTAACA 4200
Db ATATGTCAAGGCAACAGCTTACACCTTAACATCAGAACTGGGGTAAGGACCATTAACA 4200
Qy CGGGCGCTTCAATACGTAATCCACCTATGGAAGTCTCTTGGCGGACGGTGGCTGTCTG 4260
Db CGGGCGCTTCAATACGTAATCCACCTATGGAAGTCTCTTGGCGGACGGTGGCTGTCTG 4260
Qy GGGGGCTTATGACATCATATATGTGATGAGTGCCACTCACTGACTGACTGACTGACT 4320
Db GGGGGCTTATGACATCATATATGTGATGAGTGCCACTCACTGACTGACTGACTGACT 4320

QY	4321	TGGGCATCGGCACAGTCCTGGACCAAGCGGAGCGGCTGAGCGCGGCTCGTCTGCTCG	4380
Db	4321	TGGGCATCGGCACAGTCCTGGACCAAGCGGAGCGGCTGAGCGCGGCTGTCGTCTCG	4380
QY	4381	CCACCGCTACACCTCCGGGATCGTTTACCGTGCCACACCCCAATATCGAGGAATAGGCC	4440
Db	4381	CCACCGCTACACCTCCGGGATCGTTTACCGTGCCACACCCCAATATCGAGGAATAGGCC	4440
QY	4441	TGTCCAAATGAGAGATCCCTTTCTATGGCAAAAGCCATCCCAATGAGGCCATCAAG	4500
Db	4441	TGTCCAAATGAGAGATCCCTTTCTATGGCAAAAGCCATCCCAATGAGGCCATCAAG	4500
QY	4501	GGGGAGGCACTCATTTTCTGCCATTCGAAGAGAAATGTACGAGCTCGCCGAAGC	4560
Db	4501	GGGGAGGCACTCATTTTCTGCCATTCGAAGAGAAATGTACGAGCTCGCCGAAGC	4560
QY	4561	TGACAGCCTCGGACTGAACGCTGTAGCATATTACCGGGGCTTGATGTCTCGTCTATAC	4620
Db	4561	TGACAGCCTCGGACTGAACGCTGTAGCATATTACCGGGGCTTGATGTCTCGTCTATAC	4620
QY	4621	CGCCTATCGGAGACGTGTTTGTCTGGCAAAGCGCTCTAATGAAGGGTTTCAACGGCG	4680
Db	4621	CGCCTATCGGAGACGTGTTTGTCTGGCAAAGCGCTCTAATGAAGGGTTTCAACGGCG	4680
QY	4681	ATTTGACTCAGTGATCGACTGCAATACATGTGTGTCACCCAGACAGCTTCAGCTTGG	4740
Db	4681	ATTTGACTCAGTGATCGACTGCAATACATGTGTGTCACCCAGACAGCTTCAGCTTGG	4740
QY	4741	ATCCCACTTACATTTGACACAGACGCGTGGCCCCAAGACGCGTGTGCGCTCGCAAC	4800
Db	4741	ATCCCACTTACATTTGACACAGACGCGTGGCCCCAAGACGCGTGTGCGCTCGCAAC	4800
QY	4801	GGCGAGTAGAACTGGCAGGGGTAGGAGTGGCTCTACAGGTTTGTGACTCCAGGAAAC	4860
Db	4801	GGCGAGTAGAACTGGCAGGGGTAGGAGTGGCTCTACAGGTTTGTGACTCCAGGAAAC	4860
QY	4861	GGCCCTCGGGATGTTGCGATTTCTTGGTCTGTGTGAGTGTGATGACGCGGCTGTGCTT	4920
Db	4861	GGCCCTCGGGATGTTGCGATTTCTTGGTCTGTGTGAGTGTGATGACGCGGCTGTGCTT	4920
QY	4921	GGTATGAGCTCAGCCCGCTGAGACCTCGGTTAGTTGCGGCTTACTAAATACACAG	4980
Db	4921	GGTATGAGCTCAGCCCGCTGAGACCTCGGTTAGTTGCGGCTTACTAAATACACAG	4980
QY	4981	GGTTGCCCGCTGCGCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGAGCCTCACCC	5040
Db	4981	GGTTGCCCGCTGCGCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGAGCCTCACCC	5040
QY	5041	ACATAGATGCCCACTTCTGTCCAGACTAAACAGGAGGAGACAACTTTTCTTACCTGG	5100
Db	5041	ACATAGATGCCCACTTCTGTATCCAGACTAAACAGGAGGAGACAACTTTTCTTACCTGG	5100
QY	5101	TGGCATATCAAGCTACAGTGTGCCAGGCTCAAGCTCCACCTCGTGGGACCAAA	5160
Db	5101	TGGCATATCAAGCTACAGTGTGCCAGGCTCAAGCTCCACCTCGTGGGACCAAA	5160
QY	5161	TGTGGAAGTGTCTCATAGCGTGAACCTACACTGACGCGGCCAACCCCTGCTGTATA	5220
Db	5161	TGTGGAAGTGTCTCATAGCGTGAACCTACACTGACGCGGCCAACCCCTGCTGTATA	5220
QY	5221	GGCTAGGAGCGCTCCAAATGAGGTCATCTCACACACCCCAATCAATATCATATGG	5280
Db	5221	GGCTAGGAGCGCTCCAAATGAGGTCATCTCACACACCCCAATCAATATCATATGG	5280
QY	5281	CATGATGTGGGCTGACCTTGAGGTGTCATAGCACTTGGGTGTGGTAGGCGAGTCC	5340
Db	5281	CATGATGTGGGCTGACCTTGAGGTGTCATAGCACTTGGGTGTGGTAGGCGAGTCC	5340
QY	5341	TTGCAGCTTTGGCGCATACTGCTGACGACAGCAGTGTGTCTATTGTGGGAGGATCA	5400
Db	5341	TTGCAGCTTTGGCGCATATTGCTGACGACAGGAGTGTGTCTATTGTGGGAGGATCA	5400
QY	5401	TCTTGTCCGGGAAGCAGCTGTGTTCCCGACAGGGAGTCTCTTACCGAGGATTTCGATG	5460

Db	5401	TCTTGTCCGGGAAGCAGCTGTGTTCCCGACAGGGAAGTCTCTTACAGGAGTTTCGATG	5460
QY	5461	AGATGGAAGAGTGTGCTTCACAACTTCTTTAATCGAGCAGGGAATGCAAGCTCGCGAGC	5520
Db	5461	AGATGGAAGAGTGTGCTTCACAACTTCTTTAATCGAGCAGGGAATGCAAGCTCGCGAGC	5520
QY	5521	AATTCAAGCAAAAGGCGCTCGGGTTGTCGCAACCGCCACCAAGCAAGCGGAGGCTGCTG	5580
Db	5521	AATTCAAGCAAAAGGCGCTCGGGTTGTCGCAACCGCCACCAAGCAAGCGGAGGCTGCTG	5580
QY	5581	CTCCCCGTGTGGAGTCCAAAGTGGCGAGCCCTTGAGACCTTTCTGGGGAAGCAATGTGGA	5640
Db	5581	CTCCCCGTGTGGAGTCCAAAGTGGCGAGCCCTTGAGACCTTTCTGGGGAAGCAATGTGGA	5640
QY	5641	ATTTCATACGCGGAATACAGTACCTAGCAGGCTTTATCCACTCTGCTTGGGAAACCCCGCA	5700
Db	5641	ATTTCATACGCGGAATACAGTACCTAGCAGGCTTTATCCACTCTGCTTGGGAAACCCCGCA	5700
QY	5701	TAGCATCATTTGATGGCATTTACAGCTTCTATCACTAGCCGCTCACCAACCAAAACACCC	5760
Db	5701	TAGCATCATTTGATGGCATTTACAGCTTCTATCACTAGCCGCTCACCAACCAAAACACCC	5760
QY	5761	TCCTGTTTAAACATCTTGGGGGATGGTGGCTGCCAACTCTGCTCTCCAGCGCTGCGT	5820
Db	5761	TCCTGTTTAAACATCTTGGGGGATGGTGGCTGCCAACTCTGCTCTCCAGCGCTGCGT	5820
QY	5821	CAGCTTTCTGTGGCGCGGCATCCGCGAGCGGCTGTGGCAGCATAGGCTTGTGGGAAG	5880
Db	5821	CAGCTTTCTGTGGCGCGGCATCCGCGAGCGGCTGTGGCAGCATAGGCTTGTGGGAAG	5880
QY	5881	TGCTGTGGACATCTTGGCGGCTATGGGAGAGGGGTAGCGGGGCACTCGTGCCCTTTA	5940
Db	5881	TGCTGTGGACATCTTGGCGGCTATGGGAGAGGGGTAGCGGGGCACTCGTGCCCTTTA	5940
QY	5941	AGGTCAATAGCGCGAGGTGCCCTCCACCGAGACCTGTGTCAACTTACTCTCCCTGCCATCC	6000
Db	5941	AGGTCAATAGCGCGAGGTGCCCTCCACCGAGACCTGTGTCAACTTACTCTCCCTGCCATCC	6000
QY	6001	TCTCTCTGTGTGCTCGTGGGTCTGTGGCAGCAATACTCGCTCGGCACGTGG	6060
Db	6001	TCTCTCTGTGTGCTCGTGGGTCTGTGGCAGCAATACTCGCTCGGCACGTGG	6060
QY	6061	GCCCGGAGAGGGGCTGTGAGTGAACCGGCTGATAGGTTGCTTCCGCGGGGTA	6120
Db	6061	GCCCGGAGAGGGGCTGTGAGTGAACCGGCTGATAGGTTGCTTCCGCGGGGTA	6120
QY	6121	ACCAGCTCTCCCTACGCACTATGTGCTCAGAGCGAGCTGACAGCAGTGTCACTCAGA	6180
Db	6121	ACCAGCTCTCCCTACGCACTATGTGCTCAGAGCGAGCTGACAGCAGTGTCACTCAGA	6180
QY	6181	TCCTCTCTAGCCTTACCATCACTCACTGTGTAAGCGGCTCCACAGTGGATTAAAGAG	6240
Db	6181	TCCTCTCTAGCCTTACCATCACTCACTGTGTAAGCGGCTCCACAGTGGATTAAAGAG	6240
QY	6241	ACTGCTCTAGCCATGCTCGGCTGTGGCTTAAGGATGTTTGGGATTTGATATGACAG	6300
Db	6241	ACTGCTCTAGCCATGCTCGGCTGTGGCTTAAGGATGTTTGGGATTTGATATGACAG	6300
QY	6301	TGTTGACTGACTTCAAGACCTGGCTCAGTCCAACTCTTCCCGGCTTCCCGGAGTCC	6360
Db	6301	TGTTGACTGACTTCAAGACCTGGCTCAGTCCAACTCTTCCCGGCTTCCCGGAGTCC	6360
QY	6361	CTTTCTCTGTCATGCCAAACCGGGGTACAAAGGAGTCTGGCGGGGGAAGGCAATCATGCAA	6420
Db	6361	CTTTCTCTGTCATGCCAAACCGGGGTACAAAGGAGTCTGGCGGGGGAAGGCAATCATGCAA	6420
QY	6421	CCACTCTCCATCGGAGCAGATCGCGGACATGTCAAAAAAGGTTTCCATAGGATCG	6480
Db	6421	CCACTCTCCATGTGGAGCAGATCGCGGACATGTCAAAAAAGGTTTCCATAGGATCG	6480
QY	6481	TAGGCTCTAGAACCTGAGGACACAGTGGACGAAAGTTCCTCCCATCAACCATACACA	6540

Db 6481 TAGGGCTAGAACCTGCAGCAACACGTGGCATGGAACGTTCCCCATCAACGCATACACCA 6540
QY CGGACCTTGCACACCTCCCGCGCGCCAACTATTCAGGGCGCTATGGCGGTGGCTG 6600
Db 6541 CGGGACCTTGACACACCTCCCGCGCGCCAACTATTCAGGGCGCTATGGCGGTGGCTG 6600
QY CTGAGGAGTACGTGAGGAGTTAGCGGTGTGGGGGATTTCCACTATGATGACGGGCATGACCA 6660
Db 6601 CTGAGGAGTACGTGAGGAGTTAGCGGTGTGGGAGATTTCCATTAAGTACGGGATGACCA 6660
QY CTGACAAAGTAAAGTGCATGCGGATGCGGAGTTCCAGGATGCGGAGTGGGATG 6720
Db 6661 CTGACAAAGTAAAGTGCATGCGGATGCGGAGTTCCAGGATGCGGAGTGGGATG 6720
QY GAGTCGGGTGCACAGGTACGCTCCGGGTGCAGAACCTCTTCTACGGGAGGACGTCACGT 6780
Db 6721 GAGTCGGGTGCACAGGTACGCTCCGGGTGCAGAACCTCTTCTACGGGAGGACGTCACGT 6780
QY TCCAGGTCCGGGTCAACCAATATCTTGGTCCGGTCCGAGTCCCATGCGAGCCCGAACCGG 6840
Db 6781 TCCAGGTCCGGGTCAACCAATATCTTGGTCCGGTCCGAGTCCCATGCGAGCCCGAACCGG 6840
QY ACGTAACAGTGTACTTTCATGCTACCGATCCCTCCACATTAAGAGAGACGGCTA 6900
Db 6841 ACGTAACAGTGTACTTTCATGCTACCGATCCCTCCACATTAAGAGAGACGGCTA 6900
QY ACGTAAGGTGCTAGAGGGTCTCCCGCTCTTTAGCCAGCTCATCAGTACGCCAGTTGT 6960
Db 6901 ACGTAAGGTGCTAGAGGGTCTCCCGCTCTTTAGCCAGCTCATCAGTACGCCAGTTGT 6960
QY CTGCGGCTTCTTTGAAGGCGACATGACATACCCACCATGACTCCCGGACGCTGACCTCA 7020
Db 6961 CTGCGGCTTCTTTGAAGGCGACATGACATACCCACCATGACTCCCGGACGCTGACCTCA 7020
QY TCGAGGCCAACCTCTTGTGGCGGAGGAGATGGGCGGAAACATCACTCGCGTGGAGTCAG 7080
Db 7021 TCGAGGCCAACCTCTTGTGGCGGAGGAGATGGGCGGAAACATCACTCGCGTGGAGTCAG 7080
QY AGAATAAGTGTAAATCTTGGACTCTTTGAAACCGCTTACCGGAGGGGATGAGGG 7140
Db 7081 AGAATAAGTGTAAATCTTGGACTCTTTGAAACCGCTTACCGGAGGGGATGAGGG 7140
QY AGATATCCGTCCGGCGGAGATCTTCGGAATAATCCAGGAAGTTCCCTCAGCGTTGCCCA 7200
Db 7141 AGATATCCGTCCGGCGGAGATCTTCGGAATAATCCAGGAAGTTCCCTCAGCGTTGCCCA 7200
QY TATGGSCACGCCGGACTCAATCTCTACAGTCTGCTAGAGTCTTGGAAAGGACCCGGACTACG 7260
Db 7201 TATGGSCACGCCGGACTCAATCTCTACAGTCTGCTAGAGTCTTGGAAAGGACCCGGACTACG 7260
QY TCCCTCCGGTGGTACAGGATGCCATTCGCACTTACCAAGGCTCTCCAAATACCACTC 7320
Db 7261 TCCCTCCGGTGGTACAGGATGCCATTCGCACTTACCAAGGCTCTCCAAATACCACTC 7320
QY CACGAGAAAGAGGACGGTTGTCTCAGAGATFCCAATGTGTCTTCTGCTTTGGCGGAGC 7380
Db 7321 CACGAGAAAGAGGACGGTTGTCTCAGAGATFCCAATGTGTCTTCTGCTTTGGCGGAGC 7380
QY TCGCCACTAAGACCTTCCGGTACGTCGCGATCGTCCGCGTTGATAGCGACACGGGACCG 7440
Db 7381 TCGCCACTAAGACCTTCCGGTACGTCGCGATCGTCCGCGTTGATAGCGACACGGGACCG 7440
QY CCTTCTGACCTGGGCTCCGAGCAGGTGACAAAGGATCCGACGTTGAGTCGTACTCT 7500
Db 7441 CCTTCTGACCTGGGCTCCGAGCAGGTGACAAAGGATCCGACGTTGAGTCGTACTCT 7500
QY CCATGCCCTTGAAGGGGACCGCGGACCCCGATCTCAGCGACGGGTCTTGGTCTA 7560
Db 7501 CCATGCCCTTCTTGAAGGGGACCGCGGACCCCGATCTCAGCGACGGGTCTTGGTCTA 7560
QY CCGTAGTGAGGAGGTAGTGAGGATGTCTGTCTGCTGCTCAATGTCTTATACGTGGACAG 7620
Db 7561 CCGTAGTGAGGAGGTAGTGAGGATGTCTGTCTGCTGCTCAATGTCTTATACGTGGACAG 7620

QY 7621 GCGCCCTGATCAGCCCATGCGTCCGAGGAAAGTAAGCTGCCCATCAACCCGTTGAGCA 7680
Db 7621 GCGCCCTGATCAGCCCATGCGTCCGAGGAAAGTAAGCTGCCCATCAACCCGTTGAGCA 7680
QY 7681 ACTCTTTGCTGCGTCAACAACATGCTTACGCGCAACAATCCGCGAGCGCAAGCTCC 7740
Db 7681 ACTCTTTGCTGCGTCAACAACATGCTTACGCGCAACAATCCGCGAGCGCAAGCTCC 7740
QY 7741 GGCAGAGAGGTCACCTTTGACAGATTCGAGTCTGAGTGCATTTACCGGAGCGTAC 7800
Db 7741 GGCAGAGAGGTCACCTTTGACAGATTCGAGTCTGAGTGCATTTACCGGAGCGTAC 7800
QY 7801 TCAAGGAGATGAAGCGCAAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGGAG 7860
Db 7801 TCAAGGAGATGAAGCGCAAGCGCTCCACAGTTAAGGCTTAAGCTTCTATCTATAGAGGAG 7860
QY 7861 CCTCAAGCTGACGCCGCCCACTTCGGCCAAATCCAAATTTGGGTATGGGGCAAGAGCG 7920
Db 7861 CCTCAAGCTGACGCCGCCCACTTCGGCCAAATCCAAATTTGGGTATGGGGCAAGAGCG 7920
QY 7921 TCCGGAACCTATCCAGCAGGCGCGTTAACCAATCCGCTCCGCTGCGGAGGACTTGTGG 7980
Db 7921 TCCGGAACCTATCCAGCAGGCGCGTTAACCAATCCGCTCCGCTGCGGAGGACTTGTGG 7980
QY 7981 AAGACACTGAAACACCAATTCGACACCAACCATCATGGCAAAAAGTGAGGTTTTCTGCGTCC 8040
Db 7981 AAGACACTGAAACACCAATTCGACACCAACCATCATGGCAAAAAGTGAGGTTTTCTGCGTCC 8040
QY 8041 AACCAGAGAGGAGCGCGCAAGCCAGCTCGCTTATCGTATTCACAGACCTGGGAGTTC 8100
Db 8041 AACCAGAGAGGAGCGCGCAAGCCAGCTCGCTTATCGTATTCACAGACCTGGGAGTTC 8100
QY 8101 GTGTATGCGGAGAGATGSCCCTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGG 8160
Db 8101 GTGTATGCGGAGAGATGSCCCTTTACGACGTGGTCTCCACCTTCTCAGGCGGTGATGG 8160
QY 8161 GCTCTCATACGGATTTCAATACTCCCCCAAGCAGCGGTGAGTTCCTGTTGAATACT 8220
Db 8161 GCTCTCATACGGATTTCAATACTCCCCCAAGCAGCGGTGAGTTCCTGTTGAATACT 8220
QY 8221 GGAATCAAAGAAATGCCCCTATGGCTTCTCATATGACACCCGCTGTTTGACTCAACGG 8280
Db 8221 GGAATCAAAGAAATGCCCCTATGGCTTCTCATATGACACCCGCTGTTTGACTCAACGG 8280
QY 8281 TCACGTAGAGTGACATTCGTTGAGGAGTCAATTTACCAATGTTGACCTTGGCCCCCG 8340
Db 8281 TCACGTAGAGTGACATTCGTTGAGGAGTCAATTTACCAATGTTGACCTTGGCCCCCG 8340
QY 8341 AGCCAGACAGGCCATAAGTTCGCTCAGAGCGGCTTTACATCGGGGTCCCTGACTA 8400
Db 8341 AGCCAGACAGGCCATAAGTTCGCTCAGAGCGGCTTTACATCGGGGTCCCTGACTA 8400
QY 8401 ACTCAAAAAGGCGAGAACTGCGGTTATCGCGGTCCGCGCAAGTGGCGTGTGACGACTA 8460
Db 8401 ATTCAAAAAGGCGAGAACTGCGGTTATCGCGGTCCGCGCAAGTGGCGTGTGACGACTA 8460
QY 8461 GCTGCGGTAAATACCTCATGTTACTTTGAAGGCCACTGCGGCTGTGAGCTGCAAGC 8520
Db 8461 GCTGCGGTAAATACCTCATGTTACTTTGAAGGCCACTGCGGCTGTGAGCTGCAAGC 8520
QY 8521 TCCAGGACTGCACGATGCTCGTGAACGGAGCGACTTGTGTTATCTGTGAAAGCGCG 8580
Db 8521 TCCAGGACTGCACGATGCTCGTGAACGGAGCGACTTGTGTTATCTGTGAAAGCGCG 8580
QY 8581 GAAACCCAGGAGATCGCGCGCCCTACGAGTCTTCAACGAGGCTATGACTAGGTACTCCG 8640
Db 8581 GAAACCCAGGAGATCGCGCGCCCTACGAGTCTTCAACGAGGCTATGACTAGGTACTCCG 8640
QY 8641 CCCCCCGGGGATCCGCCCCCAACAGAGTACGACTCGAGCTGATACATCATGTTCT 8700
Db 8641 CCCCCCGGGGATCCGCCCCCAACAGAGTACGACTCGAGCTGATACATCATGTTCT 8700

QY	8701	CCAAATGTGTCAGTGGCGCAGATGCATCTGGGAAAAGGGTATCTACTACCTCACCCGTGACC	8760
Db	8701	CCAACTGTGTCAGTGGCGCAGATGCATCTGGGAAAAGGGTATCTACTACCTCACCCGTGACC	8760
QY	8761	CCACCAACCCCTTGCACGGCTGCGTGGGAGACAGCTAGACACACTTCCAATCAACTCTT	8820
Db	8761	CCACCAATCCCTTGCACGGCTGCGTGGGAGACAGCTAGACACACTTCCAATCAACTCTT	8820
QY	8821	GGCTAGGCAATATCATCATGTATGCGCCCAACCTTATGGGCAAGGATGATTTCTGATGACTC	8880
Db	8821	GGCTAGGCAATATCATCATGTATGCGCCCGCCCTATGGGCAAGGATGATTTCTGATGACTC	8880
QY	8881	ACTTTTCTCCATCTTCTAGCTCAAGAGCAACTTGAAGAAGCCCTGGATTCAGATCT	8940
Db	8881	ACTTTTCTCCATCTTCTAGCTCAAGAGCAACTTGAAGAAGCCCTGGATTTGTCAGATCT	8940
QY	8941	ACGGGGCTGCTACTCATTGAGCCACTTCCAGCTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9000
Db	8941	ACGGGGCTGCTACTCATTGAGCCACTTCCAGCTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9000
QY	9001	GTCTTAGCGCATTTACACTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060
Db	9001	GTCTTAGCGCATTTACTCTCCACAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9060
QY	9061	GCCTCAGGAAACTTGGGGTACCACTTCCAGCTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9120
Db	9061	GCCTCAGGAAACTTGGGGTACCACTTCCAGCTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCAT	9120
QY	9121	GCCTTAGCTACTCTCCAGGGGGGAGCGCCACTTCTGGCAGATCACTCTTTAACT	9180
Db	9121	GCCTTAGCTACTCTCCAGGGGGGAGCGCCACTTCTGGCAGATCACTCTTTAACT	9180
QY	9181	GGGCAGTAAAGACCAAGCTTAACTCACTCAATCCCGCGCGCTCCAGCTGAGCTTGT	9240
Db	9181	GGGCAGTAAAGACCAAGCTTAACTCACTCAATCCCGCGCTCCAGCTGAGCTTGT	9240
QY	9241	CTGGCTGGTTCGTCGCTGTTACAGCGGGGAGACATATATCACAGCTGCTCGTCCGC	9300
Db	9241	CGGCTGGTTCGTCGCTGTTACAGCGGGGAGACATATATCACAGCTGCTCGTCCGC	9300
QY	9301	GACCCCGCTGGTTCGCTGCTGCTACTCTTCTAGGGGTAGGCAATTTACCTGC	9360
Db	9301	GACCCCGCTGGTTCGCTGCTGCTACTCTTCTAGGGGTAGGCAATTTACCTGC	9360
QY	9361	TCCCAACCGATGAACGGGAGCTAAACCACTCCAGGCTTAAAGCCATTTCTGTTTTTTT	9420
Db	9361	TCCCAACCGATGAACGGGAGCTAAACCACTCCAGGCTTAAAGCCATTTCTGTTTTTTT	9420
QY	9421	TT	9448
Db	9421	TT	9448
RESULT 9	AF054255	8780 bp RNA linear VRL 02-MAY-1998	
LOCUS	Hepatitis C virus HC-J4, pCV-J4L6 polyprotein gene, partial cds.		
DEFINITION	Hepatitis C virus		
ACCESSION	AF054255		
VERSION	AF054255.1		
KEYWORDS			
SOURCE			
ORGANISM	Hepatitis C virus		
REFERENCE	1 (bases 1 to 8780)		
AUTHORS	Yanagi,M., St Claire,M., Shapiro,M., Emerson,S.U., Purcell,R.H. and Bukh,J.		
TITLE	Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo		
JOURNAL	Virology 244 (1), 161-172 (1998)		
PUBMED	9581788		
REFERENCE	2 (bases 1 to 8780)		
AUTHORS	Yanagi,M. and Bukh,J.		

TITLE	Direct Submission	Query Match	91.5%;	Score 8780;	DB 13;	Length 8780;
JOURNAL	Submitted (18-MAR-1998) Hepatitis Viruses Section, LID, NIAID, NIH, 7 Center Dr., Building 7, Room 201, Bethesda, MD 20892, USA	Best Local Similarity	100.0%;	Pred. No. 0;		
FEATURES	Location/Qualifiers	Matches 8780;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
source	1. .8780					
	/organism="Hepatitis C virus"					
	/mol_type="genomic RNA"					
	/strain="HC-J4"					
	/isolate="pCV-J4L6"					
	/db_xref="taxon:11103"					
	187. .>8780					
	/codon_start=1					
	/product="polyprotein"					
	/protein_id="AAC15729.1"					
	/db_xref="GI:3098648"					
CDS	/translation="MSTNPKPQRKTKRNTNRPRQDKVFGGGQIVGGVILLPRGPRRL GYATRKASERSQPRRRQPIPKARRPEGRAWAQGYPMPLFYNEGISLWAGWLLSPRG SRPSPGPTPRRRSRNLGVIDTLTCGFADLMGYIPLVAGPLGGAARLAHGVRLVED GVNYATGNLPGCSFSIFLLALLSCLTIPASAYEVNRVSGIYHVRVDLLVGTAAFCSAMV VIMHTGCVPCVQEGNSRCWALPTLAARNASVPTTIRRHVDLLVGTAAFCSAMV VGLDGSIFLVQLFTFSRPHETVQDCNSIYPGHVSHRMADMMWMSPTTALVV SOLLRIPQAVDMVAGHWGLAGLAYISWGNWAKVLLVALLFAGVDGTHHTGRVA GHTTSQFTLSFGSAGSKIQLVNTNSWHINRTALNCNDSLQTFQFAALFYAHKFNSS GCPERMASCRPIDWFAQWGPITYTKPNSDQRPYCHWYAPRCVGVVPASQVCGPVYC FTSPVVGTTDRSGVPTYSWGENETVMLNNTPRQGNWFCGTWNSTGFTKTCGG PCNIGGVNRLTICTDCERKHPATYKCGSGPMLTPRCLVDYPRVWHYPTCLNH SIFKVMYVGGVHRLNAAACNWRGRCNLEDRSELSPLLSTTTEWOLLPCAFTTL PALSTGLIHLHQNIVDVQYLYGVGSFAIKWEYILLLFLGLADARVACACUWMLL IAQEAALBNVLVNAASVAGHGILSFLVFCFAWYIKRGLAPGAAYFYGVWPLLL LIALPPLRAYALDRMAASCAGAVLGLVFLTSPYKVFLETLIWMLOYFTIRAEAH MOVVPLNAVGRGDALILLTCAVHPELIDITKLLAILGLPLMWLAQGITRYVYFVR AQLIIPACMLKRVAGGVQVFMKLGALTGTVYVYNHLTPLRDAWAGURDLAVAE PVYSWETKVTIWTGADTTAAGDIIILGLPVSARGKEIFLGPADSLGEGWRLLAPEIT AVSQRTGRKVCIIITSLTRDKNQVEGVOVSTATQTSFLATCINGCWTYHAGAGSK TLAGKPGITQMTYNTVDLVLWQAPGARSMTPCSCGSDLYLVTRHADVIYVRRRG DSRGLSPRPVSYLKGSGGILLCPSGHVGVFRAAVCTRGAKAVLTPVPMSETT MRSPVTDNSTPEAPVQTFQVAHLHAPGSKGSKTKVPAAYAGCYKVLNLPNSVAAFL GFGAYMSKAGIDPNRTGVRTITGSGIITYSTYKFLADGGSGGAYDIIICDECHS TDSITLIGTGLVDQAEATGARLVAVLATATPGSVTPVHPNIEIGLSNNGEIPYCK AIPETAKGGRHLIFCHSKKKCDELAAKTGLGLNAVAYYRGLDQSVIPIPGDVVVVA TDALMTGTFDSVIDCNTVQTFDSLDPTTETTTTTPQDASVRSRRRTGRG RSLGIRYFVTPGSRPSSVLCCEVDGAWVELTPAETSVRLARLANTPLGPVQC DHLEFVESVPTGLTHIDAFHSOTKQAGDNFVLYVAQATVCARQAAPPSPDMQMK LIRLKLTHGPTFLLYRLGAVQNEVILTHPTKYIMACMSADLEVSTWLVGSGVLA ALAAVCLTTSVIVIGRIILSGRPAPVVDREVLYQFDEMEBCAQIPIYEQGNLAE QFKQKALGLQATKQAEAPVVEKRALETFWAKHMNFISGIIQIAGLSLTPGN PATASLMAFTASITSLTQNTLLFNILGWSAAQLAPPASAFVAGALVAGVCSAA GLCKVLVDILAGYAGVAGALVAFKWSGEPSTEDLVNLLPAILSVGVVGVCA ILRRHVPGEGAVQWNNRLIFASRGNHVSPTHVVPESDAAARVTOILSSLTIQQLK RLHQWINEDCSTPCSGSLRDVNDICTVLDTFKTLQSKLLPLPGVPLSCORGK GVNRGDIMQTTCPGQAQIAGHVGSMRIVGPRCTNSNTWHGTFPINAYITGCTPSP APNYSRALMRVAAEYVEVTRVGDVHYVTGTTDNDVKCPQCPAPEFTEVDGVRLLHR YAPACKPLREDVTFOVLNOYLVGSQLPCEPEPDVTLTSMLTDPDSHITAETAKRL ARGSPSLASSASQASPLSKATCTTHDSPDADLIEANLWRQEMGNITRVESEN KVLVDSFELPHAGDEBERLSVAABILLKSKRKPSPALPIWARPDPNPLLESWKDPDY VPPVHGCPPTKAPPIPPRRKRTVLTETSNVSSALAEATKTFSSGSSAVDSGT ATAPLDASDDGKGSVDSESNMPLLEGPDPLDLSGWSVTSEASEDSEVCCSSMS YTTWGLITPCAAEESKLPINLSNLRHHNMVYATTSASRLAKKVTDFRQVLD DHYRDLVKEMKASTVAKLLSTIEEACKLTPHSAKSKFGYGAADVRLSSRAVNH RSVWEDLLETETPIDTTIMAKSEVFCVQPEKGRKPARLIVFDLGVRCVKALVD VSTLPOAVMGSSYGFQSPKQVFEVLANVTWKKKCPMGSYDTRCFDSTVETSDIRV EESIYCCDLAPAEQAIRSLERLITGGLTBLTKSKQNCQNGYRCRACGSLTSCGNTL TCVLKATPACRAAKLODCTMLVNGDLDLVICESAGTQEDAAALRAFTEAMTRYBAPPG DPPOPEYDLELITSCSSNSVAHDASGRVYLYTRDPTTPLARAAMETARTPTINSML GNITMYAPTLMWMLMTHTFSLILAQEQLEKALDC"					
QY	156	CCGCTGAGTACACCGGAATCCAGGACGACCGGGTCTTTCTTGATCAACCCGCTCAA	215			

Db 1 CCGGTGAGTACACCGGAATTGTCAGGACGACGGGTCTCTTTCTTGGATTCAACCCGCTCAA 60
QY 216 TGCCTGGAGATTGTTGGCGTGCCTCCGCGAGACTGCTAGCGAGTAGTGTGTTGGTTCGCGAA 275
Db 61 TGCCTTGGAGATTGTTGGCGTGCCTCCGCGAGACTGCTAGCGAGTAGTGTGTTGGTTCGCGAA 120
QY 276 AGGCTTGTGTGTAATGCTCTGATAGGTTGTTGCGAGTGCCTCCGCGAGTCTCTGTAAGCCG 335
Db 121 AGGCTTGTGTGTAATGCTCTGATAGGTTGTTGCGAGTGCCTCCGCGAGTCTCTGTAAGCCG 180
QY 336 TGCACATAGAGACAGAAATCTTAAACCTCAAGAAAAACCAACGTAACACCAACCGCGCG 395
Db 181 TGCACATAGAGACAGAAATCTTAAACCTCAAGAAAAACCAACGTAACACCAACCGCGCG 240
QY 396 CCACAGGAGCTCAAGTTCCCGGCGGTGCTAGATCGTTGTTGGAGTTTACTGTTGCGG 455
Db 241 CCACAGGAGCTCAAGTTCCCGGCGGTGCTAGATCGTTGTTGGAGTTTACTGTTGCGG 300
QY 456 CGCAGGGGCCAGGTTGGGTGTGCGCGACTAGGAAGGCTTCCGAGCGGTGCGAACT 515
Db 301 CGCAGGGGCCAGGTTGGGTGTGCGCGACTAGGAAGGCTTCCGAGGGTTCGCAACT 360
QY 516 CBTGAAGGCGACAACTTATCCAAAGGCTCCGACCCGAGGGCAAGGCTTGGGCTCAG 575
Db 361 CGTGAAGGCGACAACTTATCCAAAGGCTCCGACCCGAGGGCAAGGCTTGGGCTCAG 420
QY 576 CCGGGTACCTTGGCCCTCTATGGCAATGAGGCTGGGGTGGGAGGATGGCTCCTG 635
Db 421 CCGGGTACCTTGGCCCTCTATGGCAATGAGGCTGGGGTGGGAGGATGGCTCCTG 480
QY 636 TCACCCCGCGGCTCCGGCTAGTTGGGCGCCACGAGACCCCGCGTAGGTTCGCGTAAC 695
Db 481 TCACCCCGCGGCTCCGGCTAGTTGGGCGCCACGAGACCCCGCGTAGGTTCGCGTAAC 540
QY 696 TTGGGTAAAGTCAATGATACCTTATACATGCGCTTTCGCGATCTCATGGGTACATTCG 755
Db 541 TTGGGTAAAGTCAATGATACCTTATACATGCGCTTTCGCGATCTCATGGGTACATTCG 600
QY 756 CTCGTGCGCGCCCTTAGGGGCGCTGCCAGGCTTGGCACAGGTTGTCGGGTTCTG 815
Db 601 CTCGTGCGCGCCCTTAGGGGCGCTGCCAGGCTTGGCACAGGTTGTCGGGTTCTG 660
QY 816 GAGGACGGGTGAATATGACAAAGGAACTTTCGCGGTGCTCTTCTCTATCTTCCTC 875
Db 661 GAGGACGGGTGAATATGACAAAGGAACTTTCGCGGTGCTCTTCTCTATCTTCCTC 720
QY 876 TTGGCTCTGCTGCTGTTGACATCCAGCTTCGCTTATGAAGTGGCAACGTTGCC 935
Db 721 TTGGCTCTGCTGCTGTTGACATCCAGCTTCGCTTATGAAGTGGCAACGTTGCC 780
QY 936 GGGATATACATGTACGAAAGACTGCTCCAATCAAGCATTTGTATGAGGACGCGGAC 995
Db 781 GGGATATACATGTACGAAAGACTGCTCCAATCAAGCATTTGTATGAGGACGCGGAC 840
QY 996 GTGATCATCATACTCCCGGGTGGCGCTGTTGTTGAGGAGGTAACAGCTCCCGTTGC 1055
Db 841 GTGATCATCATACTCCCGGGTGGCGCTGTTGTTGAGGAGGTAACAGCTCCCGTTGC 900
QY 1056 TGGGTAGCGCTCACTCCACGCTCGCGGCAGGAATGCCAGGTCCTCCACTACGACAATA 1115
Db 901 TGGGTAGCGCTCACTCCACGCTCGCGGCAGGAATGCCAGGTCCTCCACTACGACAATA 960
QY 1116 CGACGCCAGCTGACTTGTCTGTTGGGACGGCTGCTTCTGCTCGCTATGATAGTGGG 1175
Db 961 CGACGCCAGCTGACTTGTCTGTTGGGACGGCTGCTTCTGCTCGCTATGATAGTGGG 1020
QY 1176 GATCTCTGGGATCTATTTTCTCTGCTCCAGCTGTTCACTTCTGCTCCGCTCCGCGCAT 1235
Db 1021 GATCTCTGGGATCTATTTTCTCTGCTCCAGCTGTTCACTTCTGCTCCGCTCCGCGCAT 1080
QY 1236 GAGACAGTGCAGGACTGCAACTGCTCAATCTATCCCGGCGCATGTATCAGGTCAACGCAATG 1295
Db 1081 GAGACAGTGCAGGACTGCAACTGCTCAATCTATCCCGGCGCATGTATCAGGTCAACGCAATG 1140

QY 1296 GCTTGGATATGATGAATACTGTCACCTACCAACAGCCCTAGTGTGTCGAGTTGCTC 1355
Db 1141 GCTTGGATATGATGAATACTGTCACCTACCAACAGCCCTAGTGTGTCGAGTTGCTC 1200
QY 1356 CGGATCCCAAGCTGTCGTGGACATGTCGGCGGGGGCCACCTGGGAGTCTCTGGCGGC 1415
Db 1201 CGGATCCCAAGCTGTCGTGGACATGTCGGCGGGGGCCACCTGGGAGTCTCTGGCGGC 1260
QY 1416 CTTGCCTACTATTCCATGTTAGGAACTGGGCTAAGGTTCTGATTGTGGCGCTACTCTTT 1475
Db 1261 CTTGCCTACTATTCCATGTTAGGAACTGGGCTAAGGTTCTGATTGTGGCGCTACTCTTT 1320
QY 1476 GCCCGGCTTGACGGGGAGACCCACACGACGGGGAGGTGGCGGCCACACCACTCCCGG 1535
Db 1321 GCCCGGCTTGACGGGGAGACCCACACGACGGGGAGGTGGCGGCCACACCACTCCCGG 1380
QY 1536 TTACGCTCCCTTTTCTCATCTGGGGCGTCTCAGAAAAATCCAGCTTGTGAATACCAACGCG 1595
Db 1381 TTACGCTCCCTTTTCTCATCTGGGGCGTCTCAGAAAAATCCAGCTTGTGAATACCAACGCG 1440
QY 1596 AGCTGGCACATCAACAGGACTGCCCTAAATTCGAATGCACTCCCTCCAACTGGGTTCTTT 1655
Db 1441 AGCTGGCACATCAACAGGACTGCCCTAAATTCGAATGCACTCCCTCCAACTGGGTTCTTT 1500
QY 1656 GCCCGCTGTTTACGACACAAAGTTCAAATCTGTCGGGTGCCCGGAGCGCATGGCCAGC 1715
Db 1501 GCCCGCTGTTTACGACACAAAGTTCAAATCTGTCGGGTGCCCGGAGCGCATGGCCAGC 1560
QY 1716 TGCCGCCCCATGTACTGTTGCGCCAGGGGTGGGGCCCATCACTATACCTAAGGCTTAAC 1775
Db 1561 TGCCGCCCCATGTACTGTTGCGCCAGGGGTGGGGCCCATCACTATACCTAAGGCTTAAC 1620
QY 1776 AGCTCGGATCAGAGGCTTATTGCTGCGCATTTACGCGCTCGACCGTGTGGTGTGTAACC 1835
Db 1621 AGCTCGGATCAGAGGCTTATTGCTGCGCATTTACGCGCTCGACCGTGTGGTGTGTAACC 1680
QY 1836 CGCTCGGAGGTGTCGTGTCAGTGTATTGTTTACCCCAAGCCCTGTTGTGGGGGACC 1895
Db 1681 CGCTCGGAGGTGTCGTGTCAGTGTATTGTTTACCCCAAGCCCTGTTGTGGGGGACC 1740
QY 1896 ACCGATCGTTCGGGTGTCCTACGTATAGCTGGGGGGAGAAATGAGACAGACGTGATGTC 1955
Db 1741 ACCGATCGTTCGGGTGTCCTACGTATAGCTGGGGGGAGAAATGAGACAGACGTGATGTC 1800
QY 1956 CTCAACAAACACGCTCCGCCCAAGGCAACTGTTGTCGGCTGTACATGATGAATAGTACT 2015
Db 1801 CTCAACAAACACGCTCCGCCCAAGGCAACTGTTGTCGGCTGTACATGATGAATAGTACT 1860
QY 2016 GGGTTCACTAAGACGTGCGGAGGTCCCGCTGTAACATCGGGGGGTCCGTAACCGCACC 2075
Db 1861 GGGTTCACTAAGACGTGCGGAGGTCCCGCTGTAACATCGGGGGGTCCGTAACCGCACC 1920
QY 2076 TTGATCTGCCCCACGGACTGCTTCCGGAAGCACCCCGAGGCTACTTACACAAAATGTGGC 2135
Db 1921 TTGATCTGCCCCACGGACTGCTTCCGGAAGCACCCCGAGGCTACTTACACAAAATGTGGC 1980
QY 2136 TCGGGGCCCTCGTTGAACACTAGTGTGCTAGTACCTACCATACAGGCTTTGGCACTAC 2195
Db 1981 TCGGGGCCCTCGTTGAACACTAGTGTGCTAGTACCTACCATACAGGCTTTGGCACTAC 2040
QY 2196 CCCTGCACTCTCAATTTTCCATCTTTAAGTTAGGATGTATGGGGGGGTGGAGGAC 2255
Db 2041 CCCTGCACTCTCAATTTTCCATCTTTAAGTTAGGATGTATGGGGGGGTGGAGGAC 2100
QY 2256 AGGCTCAATGCGGATGCAATTTGGACTCGAGGAGCGCTGTAACTTCGAGACAGGAT 2315
Db 2101 AGGCTCAATGCGGATGCAATTTGGACTCGAGGAGCGCTGTAACTTCGAGACAGGAT 2160
QY 2316 AGGTCAGAACTCAGCCCGCTGCTGTCTTCAACAGAGTGGCAGATACTGCCCTGTGCT 2375
Db 2161 AGGTCAGAACTCAGCCCGCTGCTGTCTTCAACAGAGTGGCAGATACTGCCCTGTGCT 2220

Qy	2376	TTCA	CCCTACCGGCTTTATCCATCGTGTGGATCCATCTCATCAGAACATCGTGAC	2435
Db	2221	TTCA	CACCCTACCGGCTTTATCCA	2271
Qy	2436	GTCA	ATACCTGTACCGGTGAGGTCAGCGTTGTCTCTTTGCAATCAAAATGGAGTAC	2495
Db	2281	GTGA	ATACCTGTACCGGTGAGGTCAGCGTTGTCTCTTTGCAATCAAAATGGAGTAC	2340
Qy	2496	ATC	TGTTGCTTTTCTTCTTCCTGGCAGACGCGCGGTGTGTGCTTCTGTGGATGATG	2555
Db	2341	ATC	TGTTGCTTTTCTTCTTCCTGGCAGACGCGCGGTGTGTGCTTCTGTGGATGATG	2400
Qy	2556	CTGC	TGATAGCCAGGCTGAGGCGGCTTATAGAACTTGGTGGTCTCTCAATGGGGCTCC	2615
Db	2401	CTGC	TGATAGCCAGGCTGAGGCGGCTTATAGAACTTGGTGGTCTCTCAATGGGGCTCC	2460
Qy	2616	GTGC	CGGAGCGATGGTATTTCTCTCTTTGTTGTTCTTCGTCGCGCGCTGGTACATT	2675
Db	2461	GTGC	CGGAGCGATGGTATTTCTCTCTTTGTTGTTCTTCGTCGCGCGCTGGTACATT	2520
Qy	2676	AAGG	CAGGCTGGCTCTGGGGGGCGTATGCTTTTTATGGCGTATGGCGCTGCTCTTG	2735
Db	2521	AAGG	CAGGCTGGCTCTGGGGGGCGTATGCTTTTTATGGCGTATGGCGCTGCTCTTG	2580
Qy	2736	CTCT	CTATGCGGTTTACAACACAGAGCTTACGCCTTGGACCGGAGATGGCTGATCGTGC	2795
Db	2581	CTCT	CTATGCGGTTTACCACACAGAGCTTACGCCTTGGACCGGAGATGGCTGATCGTGC	2640
Qy	2796	GGGG	TGCGGTTCTTGTGAGCTCTGGTATTTCTTGACCTTGTACCATCTACAAGTGT	2855
Db	2641	GGGG	TGCGGTTCTTGTGAGCTCTGGTATTTCTTGACCTTGTACCATCTACAAGTGT	2700
Qy	2856	CTCA	TAGGCTCATATGGTGTACAATCTTTATCACAGAGCGAGGCGCACATGCAA	2915
Db	2701	CTCA	TAGGCTCATATGGTGTACAATCTTTATCACAGAGCGAGGCGCACATGCAA	2760
Qy	2916	GTG	TGGTTCCTCCCTTCAACGTTTCGGGAGCGCGATGCCATCATCTCTCAAGTGT	2975
Db	2761	GTG	TGGTTCCTCCCTTCAACGTTTCGGGAGCGCGATGCCATCATCTCTCAAGTGT	2820
Qy	2976	CGGG	TTCACAGAGTTAATTTTTTGACATCACCAAACTCTCTCGGCATCTCGGCCG	3035
Db	2821	CGGG	TTCACAGAGTTAATTTTTTGACATCACCAAACTCTCTCGGCATCTCGGCCG	2880
Qy	3036	CTCA	TGTGCTCCAGGCTGGCATAAACAGAGTGCCTTCTGTCGCGCTCAAGGGCTC	3095
Db	2881	CTCA	TGTGCTCCAGGCTGGCATAAACAGAGTGCCTTCTGTCGCGCTCAAGGGCTC	2940
Qy	3096	ATT	CGTGATCATGTTAGTGCAGAAAGTCGCGGGGGTCATTATGTCAAATGGTCTTC	3155
Db	2941	ATT	CGTGATCATGTTAGTGCAGAAAGTCGCGGGGGTCATTATGTCAAATGGTCTTC	3000
Qy	3156	ATGA	AGCTGGCGGCTGACAGGTACGTATCGTTTTATACCATCTTACCCACATCGCGGAC	3215
Db	3001	ATGA	AGCTGGCGGCTGACAGGTACGTATCGTTTTATACCATCTTACCCACATCGCGGAC	3060
Qy	3216	TGGG	CCCAACGCGGCTACAGACCTTCGCGTGGCGGTAGAGCCCGCTCGTCTTCGCGC	3275
Db	3061	TGGG	CCCAACGCGGCTACAGACCTTCGCGTGGCGGTAGAGCCCGCTCGTCTTCGCGC	3120
Qy	3276	ATGA	GACCAAGGTTCATCACCTGGGGAGCAGACACCGCTGCGGTGGGGATCATCTTG	3335
Db	3121	ATGA	GACCAAGGTTCATCACCTGGGGAGCAGACACCGCTGCGGTGGGGATCATCTTG	3180
Qy	3336	GGT	CTACCGCTTCGCGCCGAAGGGGGAAGAGATATTTTGGGACCGGCTGATGCTC	3395
Db	3181	GGT	CTACCGCTTCGCGCCGAAGGGGGAAGAGATATTTTGGGACCGGCTGATGCTC	3240
Qy	3396	GAAG	GGCAGGGTGGCGACTCTCTTGCGCCCTACCGGCTACTCCCAACAAACCGCGGGC	3455
Db	3241	GAAG	GGCAGGGTGGCGACTCTCTTGCGCCCTACCGGCTACTCCCAACAAACCGCGGGC	3300
Qy	3456	GTAC	TGTGTGTCATCATCACTAGCCTCAACGCGCGGGAACAAGACCGAGGTGCAAGGGAG	3515

3301	GTACTTGGTTGGCAATCACTAGCCTTCACGGCGGGACAGAACAGAGTTCAGAGGGAG	3360
3316	GTTCAAAGTGTTTCTACCGCAACACAATCTTTCTTGGCGAAGCTGATCAACGGCGTGTGC	3375
3361	GTTCAAAGTGTTTCTACCGCAACACAATCTTTCTTGGCGAAGCTGATCAACGGCGTGTGC	3420
3376	TGCACTGTCTACCATGGCGGTGGCTCGAAGACCTTAGCGGTCCAAAGGTCCTCAATCACC	3635
3421	TGCACTGTCTACCATGGCGGTGGCTCGAAGACCTTAGCGGTCCAAAGGTCCTCAATCACC	3480
3636	CAAAATGTACCAATGTAGACCTTGACCTCGTGGCTGGCAGCGCCCCCGGGCGCGC	3695
3481	CAAAATGTACCAATGTAGACCTTGACCTCGTGGCTGGCAGCGCCCCCGGGCGCGC	3540
3696	TCCATGACACCATGCAAGCTGTGCACTCGGACCTTTTACTTGGTCAAGAGCATGCTGAT	3755
3541	TCCATGACACCATGCAAGCTGTGCACTCGGACCTTTTACTTGGTCAAGAGCATGCTGAT	3600
3756	GTCAATTCGGGTGGCGCGGAGGCGACAGACGGGAAGTCTACTCTCCCCCAGGCCGCTC	3815
3601	GTCAATTCGGGTGGCGCGGAGGCGACAGACGGGAAGTCTACTCTCCCCCAGGCCGCTC	3660
3816	TCCTACTGAAAGGCTCCTCGGTGGTCAATTGCTTTGGCTTCGGGGGACGTCGTGGGC	3875
3661	TCCTACTGAAAGGCTCCTCGGTGGTCAATTGCTTTGGCTTCGGGGGACGTCGTGGGC	3720
3876	GTCTTCCGGGCTGCTGTGTGSCACCCCGGGGGTCCGAAAGCGGTGACATTACCCGTT	3935
3721	GTCTTCCGGGCTGCTGTGTGSCACCCCGGGGGTCCGAAAGCGGTGACATTACCCGTT	3780
3936	GAGTCTATGAAATACCATGCGGTCTCCGGTCTTCACAGACAACCTCAACCCCGCGCT	3995
3781	GAGTCTATGAAATACCATGCGGTCTCCGGTCTTCACAGACAACCTCAACCCCGCGCT	3840
3996	GTACCGCAGACATTTCAAAGTGGCACATCTGCAGCTCTTACTGGCAGCGCAAGAGCACC	4055
3841	GTACCGCAGACATTTCAAAGTGGCACATCTGCAGCTCTTACTGGCAGCGCAAGAGCACC	3900
4056	AAAGTCCGGCTGCGTATGCAGGCCAAGGGTACAAGGTGCTCGTCTGAACCCGTCGTT	4115
3901	AAAGTCCGGCTGCGTATGCAGGCCAAGGGTACAAGGTGCTCGTCTGAACCCGTCGTT	3960
4116	GCGCCACCTTAGGGTTTGGGGCGTATATGTCCAAGCACACGGTATCGACCTACATC	4175
3961	GCGCCACCTTAGGGTTTGGGGCGTATATGTCCAAGCACACGGTATCGACCTACATC	4020
4176	AGAACTGGGTAAAGGACATTTACACGGGGCGGTCCAAATACGTACTCCACTATGCAAG	4235
4021	AGAACTGGGTAAAGGACATTTACACGGGGCGGTCCAAATACGTACTCCACTATGCAAG	4080
4236	TTCCCTTCGCAAGGTGGCTGTTCTGGGGGCGCTATGACATCATATATGTGATGAGTGC	4295
4081	TTCCCTTCGCAAGGTGGCTGTTCTGGGGGCGCTATGACATCATATATGTGATGAGTGC	4140
4296	CACCTCAATGACTCGACTACAATCTTGGGCATCGGCAAGTCTCTGGACCAAGCGGAGCG	4355
4141	CACCTCAATGACTCGACTACAATCTTGGGCATCGGCAAGTCTCTGGACCAAGCGGAGCG	4200
4356	GCTGGAGCGGCTCGTCTCGCCACCGCTACACTCCGGGATCGGTACCGTGCCA	4415
4201	GCTGGAGCGGCTCGTCTCGCCACCGCTACACTCCGGGATCGGTACCGTGCCA	4260
4416	CACCCCAATATCGAGGAAATAGGCGCTGTCCAACAAATGGAGATCCCGTTCTATGGCAA	4475
4261	CACCCCAATATCGAGGAAATAGGCGCTGTCCAACAAATGGAGATCCCGTTCTATGGCAA	4320
4476	GCCATCCCAATGAGGCAATCAAGGGGGGAGGCAATCTCATTTTCTGCCATTCGAAG	4535
4321	GCCATCCCAATGAGGCAATCAAGGGGGGAGGCAATCTCATTTTCTGCCATTCGAAG	4380
4536	AAATGTGACGAGCTCGCGCAAGCTGACAGGCTCGGACTGACGCTGTAGCATATTAC	4595

4381	AAATGTGACGAGCTCGCCGCAAAAGCTGACAGGCCTCGAGCTGAAACCGTGTAGCATATTAC	4444
Qy	CGGGGCTTGATGTGTCGCTCATACGCCCTTATCGGAGACGTCTGTCGTGGCAACAGAC	4655
	CGGGGCTTGATGTGTCGCTCATACGCCCTATCGGAGACGTCTGTCGTGGCAACAGAC	4500
	GCTCTAATGACGGGTTTACCGCGGATTTTGA	4715
Qy	GCTCTAATGACGGGTTTACCGCGGATTTTGA	4560
	ACCCAGACAGTGCACCTTACAGTTCGATCCACCTTACCAATTCAGACGACGACCGTGCCC	4775
	ACCCAGACAGTGCACCTTACAGTTCGATCCACCTTACCAATTCAGACGACGACCGTGCCC	4620
Qy	CAAGACGCGGTGTCGCGCTCGCAACGCGAGGTAGAACTGGCAGGGGTAGAGTGGCATC	4835
	CAAGACGCGGTGTCGCGCTCGCAACGCGAGGTAGAACTGGCAGGGGTAGAGTGGCATC	4680
	TACAGGTTTGTGACTCCAGAGAAACCGCCCTCGGGCATGTTTCAATTCCTCGTCTCTGTGT	4895
Qy	TACAGGTTTGTGACTCCAGAGAAACCGCCCTCGGGCATGTTTCAATTCCTCGTCTCTGTGT	4740
	GAGTGTATGACGCGGCTGTGCTTGGTATGAGCTCACGCCCGCTGAGACCTCGGTTAGG	4955
	GAGTGTATGACGCGGCTGTGCTTGGTATGAGCTCACGCCCGCTGAGACCTCGGTTAGG	4800
Qy	TTGCGGCTTTACCTAAATACACAGGGTTCGCCGTCTGCCAGGACCATCTGGAGTTCTCG	5015
	TTGCGGCTTTACCTAAATACACAGGGTTCGCCGTCTGCCAGGACCATCTGGAGTTCTCG	4860
	GAGAGCGCTTTCACAGGCGCTCACCCACATAGATGCCACCTCTGTGCCAGACTAAACAG	5075
Qy	GAGAGCGCTTTCACAGGCGCTCACCCACATAGATGCCACCTCTGTGCCAGACTAAACAG	4920
	GCAGGAGACAACTTTTCTTACCTGCTGGCATATCAAGCTACAGTGTGCGCCAGGGCTCAA	5135
	GCAGGAGACAACTTTTCTTACCTGCTGGCATATCAAGCTACAGTGTGCGCCAGGGCTCAA	4980
Qy	GCTCCACCTCCATCGTGGGACCAATGTGGAAGTGTCTCATAGCGCTGAAACCTACACTG	5195
	GCTCCACCTCCATCGTGGGACCAATGTGGAAGTGTCTCATAGCGCTGAAACCTACACTG	5040
	CACGGGCAACACCCCTGCTGTATAGCTTAGGAGCCGCTCCAAAATAGAGTCACTCCTCACA	5255
Qy	CACGGGCAACACCCCTGCTGTATAGCTTAGGAGCCGCTCCAAAATAGAGTCACTCCTCACA	5100
	CACCCCATAACTAAATACATCATGGCATGCAATGTCGGCTGACCTGAGGTCGTCACTAGC	5315
	CACCCCATAACTAAATACATCATGGCATGCAATGTCGGCTGACCTGAGGTCGTCACTAGC	5160
Qy	ACCTGGGTGCTGTAGCGGAGTCCCTTGCAGCTTTGGCCGCTACTGCCCTGAGACAGGC	5375
	ACCTGGGTGCTGTAGCGGAGTCCCTTGCAGCTTTGGCCGCTACTGCCCTGAGACAGGC	5220
	AGTGTGCTCATTTGTGGGAGGATCATCTTGTCCGGGAAGCAGCTGCTGTTTCCCGACAGG	5435
Qy	AGTGTGCTCATTTGTGGGAGGATCATCTTGTCCGGGAAGCAGCTGCTGTTTCCCGACAGG	5280
	GAACTCTCTACCGAGGATTCGATGAGATGGAAGAGTGTGCTCACAACCTTCCCTTACATC	5495
	GAACTCTCTACCGAGGATTCGATGAGATGGAAGAGTGTGCTCACAACCTTCCCTTACATC	5340
Qy	GAGCAGGGAATGCACTTCGCGAGCAATTCGAGCAAAAGGCGCTCGGGTTGTTGCAACG	5555
	GAGCAGGGAATGCACTTCGCGAGCAATTCGAGCAAAAGGCGCTCGGGTTGTTGCAACG	5400
	GCCACCAAGCAAGCGGAGGCTGCTGCTCCGCTGGTGGAGTCCAAGTGGCGAGCCCTTGAG	5615
Qy	GCCACCAAGCAAGCGGAGGCTGCTGCTCCGCTGGTGGAGTCCAAGTGGCGAGCCCTTGAG	5460
	ACCTTCTGGCGAAGCAATGTGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTA	5675
	ACCTTCTGGCGAAGCAATGTGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTA	5520

Qy	5676	TCACACTCTGCTTG6AAAACCCCGGATPAGCATCATTTGATGGCAATTTACAGCTTCTATCACT	5735
Db	5521	TCACACTCTGCTTG6AAAACCCCGGATPAGCATCATTTGATGGCAATTTACAGCTTCTATCACT	5580
Qy	5736	AGCCCGCTCACCACACCAAAACCCCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCC	5795
Db	5581	AGCCCGCTCACCACACCAAAACCCCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCC	5640
Qy	5796	CAACTCGCTCTCCAGCGCTGCGTCAGCTTTTCGTGGGCGCCGGCATTCGCCGAGCGGCT	5855
Db	5641	CAACTCGCTCTCCAGCGCTGCGTCAGCTTTTCGTGGGCGCCGGCATTCGCCGAGCGGCT	5700
Qy	5856	GTTGGCAGCATAGGCTTTGG6AAAGTGTCTGTGG6ACATCTTGGCGGGCTATGGGGCAGGG	5915
Db	5701	GTTGGCAGCATAGGCTTTGG6AAAGTGTCTGTGG6ACATCTTGGCGGGCTATGGGGCAGGG	5760
Qy	5916	GTAGCGGGCGCACTCGTGGCTTTTAAGTTCATGAGCGGCGAGGTGCCCTCCACCGAGGAC	5975
Db	5761	GTAGCGGGCGCACTCGTGGCTTTTAAGTTCATGAGCGGCGAGGTGCCCTCCACCGAGGAC	5820
Qy	5976	CTGCTCAACTTACTCCCTGCCATCTCTCTCTGTGTCCTCGTGTGCGGGGTCTGTGTGC	6035
Db	5821	CTGCTCAACTTACTCCCTGCCATCTCTCTCTGTGTCCTCGTGTGCGGGGTCTGTGTGC	5880
Qy	6036	GCAGCAATACTGCTGTGG6ACGTGTGGGCGCGGAGAGGGGGCTGTGCAGTGGATGAACCGG	6095
Db	5881	GCAGCAATACTGCTGTGG6ACGTGTGGGCGCGGAGAGGGGGCTGTGCAGTGGATGAACCGG	5940
Qy	6096	CTGATAGCTTTCGTTTCGGGGGTAAACAGTCTCCCTACGCACATATGTGCCTGTAGAGC	6155
Db	5941	CTGATAGCTTTCGTTTCGGGGGTAAACAGTCTCCCTACGCACATATGTGCCTGTAGAGC	6000
Qy	6156	GAGCTGCAGCAGTGTCACTCAGATCTCTCTAGCTTTACCATCACTCAACTGCTGAAG	6215
Db	6001	GAGCTGCAGCAGTGTCACTCAGATCTCTCTAGCTTTACCATCACTCAACTGCTGAAG	6060
Qy	6216	CGGCTCACCAAGTGGATTAATGAGGACTGTCTTAAGCCATGCTTCGGCTCGTGGCTAAGG	6275
Db	6061	CGGCTCACCAAGTGGATTAATGAGGACTGTCTTAAGCCATGCTTCGGCTCGTGGCTAAGG	6120
Qy	6276	GATGTTTGGGATGGGATATGCACGGTGTTCAGCTGACTTCAAGACCTTGGCTCCAGTCCAAA	6335
Db	6121	GATGTTTGGGATGGGATATGCACGGTGTTCAGCTGACTTCAAGACCTTGGCTCCAGTCCAAA	6180
Qy	6336	CTCTGCGCGGTTTACCGGGAGTCCCTTTCTGTGTCATGCCAACCGCGGTACAAAGGGATC	6395
Db	6181	CTCTGCGCGGTTTACCGGGAGTCCCTTTCTGTGTCATGCCAACCGCGGTACAAAGGGATC	6240
Qy	6396	TGGCGGGGAGCGGCATCATG3AAAACACCTGCGCCATGCGGAGCAGATGCGCGGACAT	6455
Db	6241	TGGCGGGGAGCGGCATCATG3AAAACACCTGCGCCATGCGGAGCAGATGCGCGGACAT	6300
Qy	6456	GTCAAAAACGGTTCCATGAGGATCGTAGGGCTTAGAACCTTGACACACAGTGGCACGGA	6515
Db	6301	GTCAAAAACGGTTCCATGAGGATCGTAGGGCTTAGAACCTTGACACACAGTGGCACGGA	6360
Qy	6516	ACGTTCCCCATCAACGATACACACGGGACCTTGGCACACCTCCCGCGGCCAACATAT	6575
Db	6361	ACGTTCCCCATCAACGATACACACGGGACCTTGGCACACCTCCCGCGGCCAACATAT	6420
Qy	6576	TCCAGGGGCTATGCGCGGTGGCTGTGAGGAGTACGTGAGGTTACGCTGTGGGGAT	6635
Db	6421	TCCAGGGGCTATGCGCGGTGGCTGTGAGGAGTACGTGAGGTTACGCTGTGGGGAT	6480
Qy	6636	TTCCACTAGGTGAGCGGCATGACACCATGACAAACGTAAAGTGCCCATGCCAGGTTCCGGCC	6695
Db	6481	TTCCACTAGGTGAGCGGCATGACACCATGACAAACGTAAAGTGCCCATGCCAGGTTCCGGCC	6540
Qy	6696	CCCGAAATCTTTCACGGAGGTGGATGGAGTGCGGTTGACAGGTTACGCTCGGCGGTGCAAA	6755
Db	6541	CCCGAAATCTTTCACGGAGGTGGATGGAGTGCGGTTGACAGGTTACGCTCGGCGGTGCAAA	6600

QY 6756 CCTCTTTTACGGGAGACGTCACGTTCCAGGTCGGGCTCAACCAATACCTTGTGTCGGGTCG 6815
DB 6601 CCTCTTTTACGGGAGACGTCACGTTCCAGGTCGGGCTCAACCAATACCTTGTGTCGGGTCG 6660
QY 6816 CAGCTCCCATGCGAGCCGGAACCGGACGTAACAGTCTTACTTCCATGCTCACCGATCCG 6875
DB 6661 CAGCTCCCATGCGAGCCGGAACCGGACGTAACAGTCTTACTTCCATGCTCACCGATCCG 6720
QY 6876 TCCACATATTACAGCAGAGACGGCTAAGCGTAGGCTGGCTAGAGGGTCTCCGCCCTCTTTA 6935
DB 6721 TCCACATATTACAGCAGAGACGGCTAAGCGTAGGCTGGCTAGAGGGTCTCCGCCCTCTTTA 6780
QY 6936 GCAGCTCATACGTAAGCAGTGTGCTGCGCCCTTTCTTGAAGCGCAGATGCACTACCCAC 6995
DB 6781 GCCAGCTCATACGTAAGCAGTGTGCTGCGCCCTTTCTTGAAGCGCAGATGCACTACCCAC 6840
QY 6996 CATGACTCCCCGACCGCTGACCTCATCGAGGCCCAACTCTTGTGGCGGAGAGATGGGC 7055
DB 6841 CATGACTCCCCGACCGCTGACCTCATCGAGGCCCAACTCTTGTGGCGGAGAGATGGGC 6900
QY 7056 GGAACATCATCTCGGTGGAGTCAGAGAAATAAGTAGTAATTCTGGACTCTTTTGGAAACCG 7115
DB 6901 GGAACATCATCTCGGTGGAGTCAGAGAAATAAGTAGTAATTCTGGACTCTTTTGGAAACCG 6960
QY 7116 CTTTACGCGGAGGGGATGAGAGGAGATATCCGTGCGCGCGAGATCTTGGGAAATCC 7175
DB 6961 CTTTACGCGGAGGGGATGAGAGGAGATATCCGTGCGCGCGAGATCTTGGGAAATCC 7020
QY 7176 AGGAAGTTCCCTCAGCGTTGCCATATGGGCAGCCCGGACTACAACTCTCCACTGCTA 7235
DB 7021 AGGAAGTTCCCTCAGCGTTGCCATATGGGCAGCCCGGACTACAACTCTCCACTGCTA 7080
QY 7236 GAGTCCCTGGAAGGACCCGGACTAGTCCCTCGGTGGTACACGGATGCCATTTGCCACCT 7295
DB 7081 GAGTCCCTGGAAGGACCCGGACTAGTCCCTCGGTGGTACACGGATGCCATTTGCCACCT 7140
QY 7296 ACCAAGGCTCCTCCAATACCACTCCACGGAGAAAGGACGGTTGTCTTGACAGAAATCC 7355
DB 7141 ACCAAGGCTCCTCCAATACCACTCCACGGAGAAAGGACGGTTGTCTTGACAGAAATCC 7200
QY 7356 AATGTGTCTTCTGCTTTGGCGGAGCTCGCACTAAGACCTTTCGGTAGCTCGGATCTGTCG 7415
DB 7201 AATGTGTCTTCTGCTTTGGCGGAGCTCGCACTAAGACCTTTCGGTAGCTCGGATCTGTCG 7260
QY 7416 GCGTTGATAGCGGACCGGACCGCCCTTCTGACCTGCGCTCCGACGAGCGTGAACA 7475
DB 7261 GCGTTGATAGCGGACCGGACCGCCCTTCTGACCTGCGCTCCGACGAGCGTGAACA 7320
QY 7476 GGATCCGACGTTGAGTCTGTAATCTCTCATGCCCCCTTGAAGGGGAGCCGGGGACCCC 7535
DB 7321 GGATCCGACGTTGAGTCTGTAATCTCTCATGCCCCCTTGAAGGGGAGCCGGGGACCCC 7380
QY 7536 GATCTCAGCGACGGGCTTGTGCTTACCGTAGTGAGGAGGCTAGTGAGGATGTCTGTCG 7595
DB 7381 GATCTCAGCGACGGGCTTGTGCTTACCGTAGTGAGGAGGCTAGTGAGGATGTCTGTCG 7440
QY 7596 TGCTCAATGTCTTATAGTGACAGGCGCCCTGATACGCCCATGCGGTCGGAGGAAAGT 7655
DB 7441 TGCTCAATGTCTTATAGTGACAGGCGCCCTGATACGCCCATGCGGTCGGAGGAAAGT 7500
QY 7656 AAGTGGCCCATCAACCCGTTGAGCAACTCTTGTGCTGCTACCAACAACATGCTTACGCC 7715
DB 7501 AAGTGGCCCATCAACCCGTTGAGCAACTCTTGTGCTGCTACCAACAACATGCTTACGCC 7560
QY 7716 ACAACATCCCGAGCGCAAGCTCCGCGAGAAAGGTCACCTTTTGAACAGATTGCAAGTC 7775
DB 7561 ACAACATCCCGAGCGCAAGCTCCGCGAGAAAGGTCACCTTTTGAACAGATTGCAAGTC 7620
QY 7776 CTGGATGATATTACCGGACGTAATCAAGGAGATGAAGGGCAAGGCTTCAAGTTAAG 7835
DB 7621 CTGGATGATATTACCGGACGTAATCAAGGAGATGAAGGGCAAGGCTTCAAGTTAAG 7680
QY 7836 GCTAAGCTTCTATATTAGAGAGCGCTGCAAGCTGACGCCCCCACTTCGGCCAAATCC 7895

DB 7681 GCTAAGCTTCTATATTAGAGAGGCTGCAAGCTGACGCCCCCACTTTCGGCCAAATCC 7740
QY 7896 AAATTTGGCTATGGGCAAAAGGACGTCGGAACCTATCCAGCAGGCGCCGTTAAACCAATC 7955
DB 7741 AAATTTGGCTATGGGCAAAAGGACGTCGGAACCTATCCAGCAGGCGCCGTTAAACCAATC 7800
QY 7956 CGCTCGGTGGGAGACTTGTGGAAGACACTGGAACCACTGAACCACTTGAACCACTCATG 8015
DB 7801 CGCTCGGTGGGAGACTTGTGGAAGACACTGGAACCACTTGAACCACTCATG 7860
QY 8016 GCAAAAGTGAGTCTTCTGCGTCCAAACAGAGAAGGGCGCGCAAGCCAGCTCGCCTT 8075
DB 7861 GCAAAAGTGAGTCTTCTGCGTCCAAACAGAGAAGGGCGCGCAAGCCAGCTCGCCTT 7920
QY 8076 ATCGTATTTCCAGACCTGGGAGTTCGTATATGCGAAGATGCCCTTTACGACGTGGTC 8135
DB 7921 ATCGTATTTCCAGACCTGGGAGTTCGTATATGCGAAGATGCCCTTTACGACGTGGTC 7980
QY 8136 TCCACCTTCTCAGGCGGTGATGGGCTCTCATAGGATTTCAATACTCCCCCAAGCAG 8195
DB 7981 TCCACCTTCTCAGGCGGTGATGGGCTCTCATAGGATTTCAATACTCCCCCAAGCAG 8040
QY 8196 CGGTCGAGTTCCTGCTGAATACCTCGAAATCAAGAAATGCCCTATGGGCTTCTCATAT 8255
DB 8041 CGGTCGAGTTCCTGCTGAATACCTCGAAATCAAGAAATGCCCTATGGGCTTCTCATAT 8100
QY 8256 GACACCCGCTGTTTGACTCAACCGTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATT 8315
DB 8101 GACACCCGCTGTTTGACTCAACCGTCACTGAGAGTGACATTCGTGTTGAGGAGTCAATT 8160
QY 8316 TACCAATGTTGTGACTTGGCCCCCGAGGCCAGACAGGCCATAGGTGCTCAGAGAGCG 8375
DB 8161 TACCAATGTTGTGACTTGGCCCCCGAGGCCAGACAGGCCATAGGTGCTCAGAGAGCG 8220
QY 8376 CTTTACATCGGGGTCCTTGACTAACTCAAAAGGCGAGAACTGCGGTTATCGCCGTCG 8435
DB 8221 CTTTACATCGGGGTCCTTGACTAACTCAAAAGGCGAGAACTGCGGTTATCGCCGTCG 8280
QY 8436 CGCGAAGTGGCTGTGACGACTAGTGGCGGTAATACCTCACATGTTACTTTGAAGGCC 8495
DB 8281 CGCGAAGTGGCTGTGACGACTAGTGGCGGTAATACCTCACATGTTACTTTGAAGGCC 8340
QY 8496 ACTGACGCTGTGAGCTGCAAGCTCCAGGACTGACAGTGTCTGTGAACGGAGAGAC 8555
DB 8341 ACTGACGCTGTGAGCTGCAAGCTCCAGGACTGACAGTGTCTGTGAACGGAGAGAC 8400
QY 8556 CTTGTCTTATCTGTGAAGCGCGGAAACCCAGAGAGATGGGCGGCCCTACGAGCCTTC 8615
DB 8401 CTTGTCTTATCTGTGAAGCGCGGAAACCCAGAGAGATGGGCGGCCCTACGAGCCTTC 8460
QY 8616 ACGGAGCTATGACTAGGTATTCGCCGCCCGGGGATCGCCCAACCAAGAAATACGAC 8675
DB 8461 ACGGAGCTATGACTAGGTATTCGCCGCCCGGGGATCGCCCAACCAAGAAATACGAC 8520
QY 8676 CTGAGCTGATAACATCATGTTCTCCAATGTGTGTCAGTCGCGCAGCATGCTTGGCAAA 8735
DB 8521 CTGAGCTGATAACATCATGTTCTCCAATGTGTGTCAGTCGCGCAGCATGCTTGGCAAA 8580
QY 8736 AGGTATATACCTCACCCGTGACCCCAACCCCTTTGACGCGGTGCGTGGAGACA 8795
DB 8581 AGGTATATACCTCACCCGTGACCCCAACCCCTTTGACGCGGTGCGTGGAGACA 8640
QY 8796 GCTAGACACTCAATCAACTCTTGGCTAGGCAATATCATCATGATGCGCCCAACCTTA 8855
DB 8641 GCTAGACACTCAATCAACTCTTGGCTAGGCAATATCATCATGATGCGCCCAACCTTA 8700
QY 8856 TGGCAAGGATGATTTCTGATGACTCACTTTTCTCCATCTTCTAGCTCAAGAGCAACTT 8915
DB 8701 TGGCAAGGATGATTTCTGATGACTCACTTTTCTCCATCTTCTAGCTCAAGAGCAACTT 8760
QY 8916 GAAAAAGCCCTGGATTGTCA 8935

Db	8761	CAAAAAGCCCTGGATTGCA	8780	
RESULT 10				
AF054257				
LOCUS	AF054257	8780 bp	RNA	linear
DEFINITION	Hepatitis C virus HC-J4, pcv-J4L8 polyprotein gene, partial cds.			VRL 02-MAY-1998
ACCESSION	AF054257			
VERSION	AF054257.1	GI:3098650		
KEYWORDS				
SOURCE	Hepatitis C virus			
ORGANISM	Hepatitis C virus			
REFERENCE	1 (bases 1 to 8780)			
AUTHORS	Yanagi,M., St Claire,M., Shapiro,M., Emerson,S.U., Purcell,R.H. and Bukh,J.			
TITLE	Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo			
JOURNAL	Virology 244 (1), 161-172 (1998)			
PUBMED	9581788			
REFERENCE	2 (bases 1 to 8780)			
AUTHORS	Yanagi,M. and Bukh,J.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-MAR-1998) Hepatitis Viruses Section, LID, NIAID, NIH, 7 Center Dr., Building 7, Room 201, Bethesda, MD 20892, USA			
FEATURES	Location/Qualifiers			
source	1..8780			
	/organism="Hepatitis C virus"			
	/mol_type="genomic RNA"			
	/strain="HC-J4"			
	/isolate="pcv-J4L8"			
	/db_xref="taxon:11103"			
	187..>8780			
	/codon_start=1			
	/product="polyprotein"			
	/protein_id="AAC15730.1"			
	/db_xref="GI:3098651"			
	/translation="MSTNPKPQKTKRNTNRRPQDKPFGGQIVGVYPLPRGPRL GVRTRKASERQPRKQPIPKARRPGRWAQPGYDPLVYGNELGAGWLLSPRG SRPSGPTDPRSRNLGKIDLTGCFADLMGYIPLVGAPLGAARALAHGVLEAD GVNAYTGLPCGSGSIFILALISCLTIPASAYEVNRVSGIYHVTNDSNSIYVBAAD VIMHTPGCVFQVQEGNSRCWALPTLAARNASVPTTIRRHVDLLVGTAAFCSAMV VGDLCGSIFLVSLQFTFPRHETVQDCNCSLYPGHVSCHMAWMNMWNSPTTLV SOLLIPQAVDNDVAGAHGVLGAYYSWGNWAKVILVALLPAGVGEHTHTRVV GHTTSGFTLSSGSAQKIQLVNTYNGSHINRALTNCNDSIQGTQFFALFYAHKFNSS GCPERMANCRPIDWAQOGPIYTKPNSSDQRPYCHYAPRPGVFPASQVCGPVYC FTPSVAVVNTDRSGPTYSGENETDVMNLNTRPPQGNWFGCTWMNSTGTFTKCGG PPCNTGGVGNRTLIPTDCFRKHPEATYTKCGSGPWLTPRCLVDYPIYRLMHPYCTLNF SIFKVRMYGVGHEHLNACNWRGECNLEDRSELSPILLSTTEWPIILPCAFTLL PALSTGLHLHONIYDVQVLYGVGSAFVSFAIKWEYIILLFLLLADARVCACLWMMLL IAQAABALENVVLNAAVSAGAHGILSLVFFCAAWYINGRLAPAGAAFYGVWPLLL LLALLPFRAYALDREMAASCGAVLGLVFLTSLPYKVFYFLRLIWLQYFITRAEAB MQWVPLNVRGDRDAIILLTCAVHPELIFDITKLIALILGLMVLQAGITRVPYFVR AQLIRACMLKVTAGAGVHIOAMKLGALGTGVYVNHILTPLRDWAHAGRLDLAVE PVFSGMETHKVTWGDRTAACGDIILGLPVSARKEIFLGPADSLEGGWFLAPIT AISOQTRGVLCVITSLTRDKQVEGEVQVNSTATQSFPLATCINGVCWTVYHAGSK TLIAGPKGPIITQMYTNDVLDGVQAPPGARMTPCSGSSDLVLRHADVLPVRRRG DLRGSLLSPRPSYILKSGSGGLLCPSGHVGVFRAACVTRGVAKAVDFIPVESMETT MRSPVTDNSPPAVPQTPQVAHLHAPTSGSKSTKVPAAVQAQGVKLVLPNSVAATL GFGVMSKAHGDIPNIRTVRTITGSGITYSTYKGFADAGCGSGAVDIIICDRCHS TDSITLIGTLVDQATAGARVLVATATPQSVTVVPHNTEELGLSNNGEIPVVGK AIPERAKGGRHLIFCHSKKCKDELAKLTLGLNLNAVAYYRGLDVSVIPPGDVVVVA TDALMTGFTGDFSDVDICNTVQTVDFSLDPTFTTETTTTTPQDAVRSQRRTGRG RSGIYRFPVTPRSGMFDSSVLCQYDAGCAWYELTPAETSVRILRAYLNTPLGDPVQC DHLEWESVFTGLTHIDHDFLSQTKQAGNFPYVAYQATVCARQAQAPPSPDQMWKC LIRLKPTLHGPTLRYLCAVONELTHPIITKYIMACMSADLEVTVSTWLVGVGLAE ALNAYCLTGVSVIVGRIIILSGKPAVVPDREVLVQEFDEMERSCASQLPVIEQGMOLAE QFKQALGLQATQKQAEAAAPVSESKRALETFNAXHWNFIISGLIYLAGLSTLPGN PAIASLMAFTASTISPLTTHQTLFNILNGWVAQAPPSAASAFVAGIAGAAVCSII GLGKLVLDIYLAGVAGALVAFKVMGSVEFSTEDLVNLLPAILSPGALVVGVCAA ILRRHVGPGEVQVMNRIIFASRGNHVSPTHYVSPESDAAARVTQILSLSTITQLLK RLHQWINECDSTPCSGSLRDLVDWMDICTVLTDFTKWLQSKLLPRLPGVPFLSCQRGYK			
CDS				

Db 2941 ATTCTGTCATGTTAGTGCAGAAAGTCGCCGGGGTCAATATATCAAAATGAGCCCTTC 3000
Qy 3156 ATGAAGCTGGGGCGCTGACAGGTACGTACGTTTATAACCACTTTACCCCACTGCGGGAC 3215
Db 3001 ATGAAGCTGGGGCGCTGACAGGTACGTACGTTTATAACCACTTTACCCCACTGCGGGAC 3060
Qy 3216 TGGGCCACCGCGGGCTTACGAGACCTTCGCGTGGCGGTAGAGCCGTCGTCTTCGCCCC 3275
Db 3061 TGGGCCACCGCGGGCTTACGAGACCTTCGCGTGGCGGTAGAGCCGTCGTCTTCGCCCC 3120
Qy 3276 ATGGAGACCAAGGTCATCACCTGGGGAGCAGACACCGCTCGGTGGGGACATCATCTTG 3335
Db 3121 ATGGAGACCAAGGTCATCACCTGGGGAGCAGACACCGCTCGGTGGGGACATCATCTTG 3180
Qy 3336 GGTCTACCCGCTCTCGCCCCGAAGGGAGAGATATTTTGGGACCGGCTGATAGTCTC 3395
Db 3181 GGTCTACCCGCTCTCGCCCCGAAGGGAGAGAGATATTTTGGGACCGGCTGATAGTCTC 3240
Qy 3396 GAAGGCAAGGTGGGACTCTCTTGGCCCATCACGGGCTACTCCCAACAAACGCGGGGC 3455
Db 3241 GAAGGCAAGGTGGGACTCTCTTGGCCCATCACGGGCTACTCCCAACAAACGCGGGGC 3300
Qy 3456 GTACTTGGTTGATCATCATCTAGCTCAGAGCCGGGACAAAGACAGGTGAAAGGGAG 3515
Db 3301 GTACTTGGTTGATCATCATCTAGCTCAGAGCCGGGACAAAGACAGGTGAAAGGGAG 3360
Qy 3516 GTTCAAGTGGTTTACCGCAACACAACTTTCTTGGGACCTGCATCAAGCGGCTGTGC 3575
Db 3361 GTTCAAGTGGTTTACCGCAACACAACTTTCTTGGGACCTGCATCAAGCGGCTGTGC 3420
Qy 3576 TGGACTGTCTACATGGCGCTGGCTCGAAGACCTTAGCGGCTCCAAAAGGTCCAATCAC 3635
Db 3421 TGGACTGTCTACATGGCGCTGGCTCGAAGACCTTAGCGGCTCCAAAAGGTCCAATCAC 3480
Qy 3636 CAAATGTACCAATGTAGACCTGAGACCTCGTGGCTGCGAGCGGCCCGCGGGCGGC 3695
Db 3481 CAAATGTACCAATGTAGACCTGAGACCTCGTGGCTGCGAGCGGCCCGCGGGCGGC 3540
Qy 3696 TCCATGACACCATGAGCTGTGGAGCTCGGACCTTTACTTGGTCAGGAGACATGCTGAT 3755
Db 3541 TCCATGACACCATGAGCTGTGGAGCTCGGACCTTTACTTGGTCAGGAGACATGCTGAT 3600
Qy 3756 GTCAATTCGGTGGCGCGGAGCGGACAGAGGGGAAGTCTACTCTCCCCCAAGGCCCGCT 3815
Db 3601 GTCAATTCGGTGGCGCGGAGCGGACAGAGGGGAAGTCTACTCTCCCCCAAGGCCCGCT 3660
Qy 3816 TCCATCTGAAAGCTCTCGGGTGTGCATTTGCTTGGCCCTTGGGGGACAGTCTGGGC 3875
Db 3661 TCCATCTGAAAGGCTCTCTCGGGTGTGCATTTGCTTGGCCCTTGGGGGACAGTCTGGGC 3720
Qy 3876 GTCTTCGGGCTGCTGTGTGCACCGGGGGGTGCGGAAGCGGTGGACTTCATACCCGTT 3935
Db 3721 GTCTTCGGGCTGCTGTGTGCACCGGGGGGTGCGGAAGCGGTGGACTTCATACCCGTT 3780
Qy 3936 GAGTCTATGGAACTACCAATGGGCTCTCGGGTCTTACAGAGCAACTCAACCCCGCGGCT 3995
Db 3781 GAGTCTATGGAACTACCAATGGGCTCTCGGGTCTTACAGAGCAACTCAACCCCGCGGCT 3840
Qy 3996 GTACCGCAGACATTCAGATGGGCATCTGCGAGCTCTTACTGTGGCAGCGGCAAGAGACCC 4055
Db 3841 GTACCGCAGACATTCAGATGGGCATCTGCGAGCTCTTACTGTGGCAGCGGTGAAGACAT 3900
Qy 4056 AAAGTCCCGGCTGCGTATCGACCCCAAGGATACAGGTGCTCGTCTGAACCGTCCGTT 4115
Db 3901 AAAGTCCCGGCTGCGTATGAGCCCAAGGATACAGGTGCTCGTCTGAACCGTCCGTT 3960
Qy 4116 GCGCCACCTTAGGGTTTGGGGCGTATATGTCCAAGGCAACGGTATCGACCCCTAACATC 4175
Db 3961 GCGCCACCTTAGGGTTTGGGGCGTATATGTCCAAGGCAACGGTATCGACCCCTAACATC 4020
Qy 4176 AGAATGGGGTAAGGACCAATTACACGCGGGCTCCATTTACGTATCTCCACCTATGGCAAG 4235
Db 4021 AGAATGGGGTAAGGACCAATTACACGCGGGCTCCATTTACGTATCTCCACCTATGGCAAG 4080

Qy 4236 TTCTTGGCCGACGGTGGCTGTTCTGGGGGCGCCTTATGACATCATATATGTGATGATGC 4295
Db 4081 TTCTTGGCCGACGGTGGCTGTTCTGGGGGCGCCTTATGACATCATATATGTGATGATGC 4140
Qy 4296 CACTCAACTGACTCGACTACCATCTTGGGCATCGGCACAGTCTCGGACCAAGCGGAGACG 4355
Db 4141 CACTCAACTGACTCGACTACCATCTTGGGCATCGGCACAGTCTCGGACCAAGCGGAGACG 4200
Qy 4356 GCTGGAGCGGGCTGCTGCTCGCCACCGCTACACCTCGGGATCGGTTACGTCGCA 4415
Db 4201 GCTGGAGCGGGCTGCTGCTCGCCACCGCTACACCTCGGGATCGGTTACGTCGCA 4260
Qy 4416 CACCCCAATATCGAGGAAATAGGCTGTCCAAATGAGAGATCCCTTCTATGGCAAA 4475
Db 4261 CACCCCAATATCGAGGAAATAGGCTGTCCAAATGAGAGATCCCTTCTATGGCAAA 4320
Qy 4476 GCGATCCCATATGAGGCCATCAAGGGGGGAGGATCTCATTTCTTGCATTTCCAAAG 4535
Db 4321 GCGATCCCATATGAGGCCATCAAGGGGGGAGGATCTCATTTCTTGCATTTCCAAAG 4380
Qy 4536 AAATGTGACGAGCTCGCGCAAAAGCTGACAGGCCTCGGACCTGAACGCTGTAGCATATTAC 4595
Db 4381 AAATGTGACGAGCTCGCGCAAAAGCTGACAGGCCTCGGACCTGAACGCTGTAGCATATTAC 4440
Qy 4596 CCGGGCTTGTATGTGTCCGTCAATACCGCTATCGGAGACGTCGTTGTCTGGCAACAGAC 4655
Db 4441 CCGGGCTTGTATGTGTCCGTCAATACCGCTATCGGAGACGTCGTTGTCTGGCAACAGAC 4500
Qy 4656 GCTCTAATGACGGGTTTACCGGCGATTTTGACTCACTGATGATCGACTGCAATACATGTGC 4715
Db 4501 GCTCTAATGACGGGTTTACCGGCGATTTTGACTCACTGATGATCGACTGCAATACATGTGC 4560
Qy 4716 ACCCAGACAGTCGACTTCAGCTTCGATCCCACTTACCATTGAGACGACGCTGCGCC 4775
Db 4561 ACCCAGACAGTCGACTTCAGCTTCGATCCCACTTACCATTGAGACGACGCTGCGCC 4620
Qy 4776 CAAGACGCGGTGTCCGCTCGCAACCGGAGGTAGAACTGGCAGGGGTAGGAGTGGCATC 4835
Db 4621 CAAGACGCGGTGTCCGCTCGCAACCGGAGGTAGAACTGGCAGGGGTAGGAGTGGCATC 4680
Qy 4836 TACAGGTTTGTGATCTCAGAGGAAACGCGCTCGGGCATGTTTCGATCTTTCGCTCTGTGT 4895
Db 4681 TACAGGTTTGTGATCTCAGAGGAAACGCGCTCGGGCATGTTTCGATCTTTCGCTCTGTGT 4740
Qy 4896 GAGTCTATGACGGGGCTGCTTGGTATGAGTCAACGCGCTGAGACCTCGGTTAGG 4955
Db 4741 GAGTCTATGACGGGGCTGCTTGGTATGAGTCAACGCGCTGAGACCTCGGTTAGG 4800
Qy 4956 TTGCGGGCTTACCTAAATACACAGGGTTGCGCGTCTGCCAGGACCATCTGGAGTTCGCG 5015
Db 4801 TTGCGGGCTTACCTAAATACACAGGGTTGCGCGTCTGCCAGGACCATCTGGAGTTCGCG 4860
Qy 5016 GAGAGCTTTCACAGGCTCACCCACATAGATGCCACTTCTGTGTCAGACTAAACAG 5075
Db 4861 GAGAGCTTTCACAGGCTCACCCACATAGATGCCACTTCTGTGTCAGACTAAACAG 4920
Qy 5076 GCAGGAGCAACTTTCCTTACTGTTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAA 5135
Db 4921 GCAGGAGCAACTTTCCTTACTGTTGGCATATCAAGCTACAGTGTGCGCAGGGCTCAA 4980
Qy 5136 GCTCCACCTCCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTGAAACCTACACTG 5195
Db 4981 GCTCCACCTCCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTGAAACCTACACTG 5040
Qy 5196 CACGGGCCAACCCCTGCTGTATAGGCTAGAGCCGCTCCAAAATGAGGTCACTCTACA 5255
Db 5041 CACGGGCCAACCCCTGCTGTATAGGCTAGAGCCGCTCCAAAATGAGGTCACTCTACA 5100
Qy 5256 CACCCCATTAATTAATACATCATGGCATGTGCGGCTGACCTGAGGTGCTCACTAGC 5315
Db 5101 CACCCCATTAATTAATTAATACATCATGGCATGTGCGGCTGACCTGGAGGTGCTCACTAGC 5160

QY	5316	ACCTGGGTGCTGTAGCGGAGTCTTTGACAGCTTTGGCGCATACTGCTGACGACGCG	5375
Db	5161	ACCTGGGTGCTGTAGCGGAGTCTTTGACAGCTTTGGCGCATACTGCTGACGACGCG	5220
QY	5376	AGTGTGGTCAATGTCGGGAGGATCATCTGTGTCGGGAAGCCAGCTGCTGCTCCGACAGG	5435
Db	5221	AGTGTGGTCAATGTCGGGAGGATCATCTGTGTCGGGAAGCCAGCTGCTGCTCCGACAGG	5280
QY	5436	GAAGTCTCTACACGAGGTTGATGAGATGGAAGAGTGTGCTTCAACAATCTCTTACATC	5495
Db	5281	GAAGTCTCTACACGAGGTTGATGAGATGGAAGAGTGTGCTTCAACAATCTCTTACATC	5340
QY	5496	GAGGAGGAATGACGTGCGGAGCAATTCAGCAAAAGCGCTCGGTGTTGTCAAACG	5555
Db	5341	GAGCAGGAATGACGTGCGGAGCAATTCAGCAAAAGCGCTCGGTGTTGTCAAACG	5400
QY	5556	GCCACCAAGCAAGCGGAGGCTGCTCCGTCGTGGAGTCCAAAGTGGCGAGCCCTTGAG	5615
Db	5401	GCCACCAAGCAAGCGGAGGCTGCTCCGTCGTGGAGTCCAAAGTGGCGAGCCCTTGAG	5460
QY	5616	ACCTTCTGGCGGAGACATGTGGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTA	5675
Db	5461	ACCTTCTGGCGGAGACATGTGGAAATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTA	5520
QY	5676	TCCACTCTGCTGGAACCCCGGATAGCATCATTTGATGGCATTTACAGCTTCTATCACT	5735
Db	5521	TCCACTCTGCTGGAACCCCGGATAGCATCATTTGATGGCATTTACAGCTTCTATCACT	5580
QY	5736	AGCCCGCTACACCAACCAACACCCCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCC	5795
Db	5581	AGCCCGCTACACCAACCAACACCCCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCC	5640
QY	5796	CAACTCTCTCTCCAGCGTGGTCACTTCTGTCGGCGCGGCATCGCGGAGCGGCT	5855
Db	5641	CAACTCTCTCTCCAGCGTGGTCACTTCTGTCGGCGCGGCATCGCGGAGCGGCT	5700
QY	5856	GTTGGCAGCATAGGCTTTGGGAAGTGTCTGTGGACATCTTGGCGGCTATGCGGCGAGG	5915
Db	5701	GTTGGCAGCATAGGCTTTGGGAAGTGTCTGTGGACATCTTGGCGGCTATGCGGCGAGG	5760
QY	5916	GTAGCCGGCGCATCTGTGGCTTTAAGTCAATGAGCGCGGAGGTGCCCTCCACCGAGAC	5975
Db	5761	GTAGCCGGCGCATCTGTGGCTTTAAGTCAATGAGCGCGGAGGTGCCCTCCACCGAGAC	5820
QY	5976	CTGTCTCACTTACTCCCTGCCATCCTCTCTCTGTCCTGTCCTGTCCTGCGGGTCTGTGC	6035
Db	5821	CTGTCTCACTTACTCCCTGCCATCCTCTCTCTGTCCTGTCCTGTCCTGCGGGTCTGTGC	5880
QY	6036	GCAGCAATACCTGCTGCGACGTGGCGCGGAGAGGGGCTGTGCAGTGGATGAACCGG	6095
Db	5881	GCAGCAATACCTGCTGCGACGTGGCGCGGAGAGGGGCTGTGCAGTGGATGAACCGG	5940
QY	6096	CTGATAGCGTTTCGTTTCGGGGTAACACAGTCTCTCCCTACGCAATATGTCCCTGAGAGC	6155
Db	5941	CTGATAGCGTTTCGTTTCGGGGTAACACAGTCTCTCCCTACGCAATATGTCCCTGAGAGC	6000
QY	6156	GACGCTCAGCAGCGTGTCACTAGATCCTCTCTAGCCTTACCATCACTCACTGCTGAAG	6215
Db	6001	GACGCTCAGCAGCGTGTCACTAGATCCTCTCTAGCCTTACCATCACTCACTGCTGAAG	6060
QY	6216	CGGCTCCACCAAGTGAATTAAGAGACGTCTCTAGCCATGCTCCGGCTCGTGGCTAAGG	6275
Db	6061	CGGCTTCACCAAGTGAATTAAGAGACGTCTCTAGCCATGCTCCGGCTCGTGGCTAAGG	6120
QY	6276	GATGTTTGGGATTTGATATGACGCGTGTGACTGACTTCAAGACCTGGCTCCAGTCCAAA	6335
Db	6121	GATGTTTGGGATTTGATATGACGCGTGTGACTGACTTCAAGACTTGGCTCCAGTCCAAAG	6180
QY	6336	CTCCTGCGCGGTTACCGGAGTCCCTTTCTGTCATGCCAACGCGGTAACAGGAGTGC	6395
Db	6181	CTCCTGCGCGGTTACCGGAGTCCCTTTCTGTCATGCCAACGCGGTAACAGGAGTGC	6240
QY	6396	TGGCGGGGGGACGGCATCATGCAAAACCACTGCCCATGCGGAGCACAGATCGCCGACAT	6455

Db	6241	TGGCGGGGGACGGCATCATGCAAAACCACTGCCCATGTGGAGCACAGATCGCCGACAT	6300
QY	6456	GTCAAAAAACGGTTTCCATGAGGATTCGTAGGGCTTAGAACTTGAGCAACACGTGGGACGGA	6515
Db	6301	GTCAAAAAACGGTTTCCATGAGGATTCGTAGGGCTTAGAACTTGAGCAACACGTGGGACGGA	6360
QY	6516	AGTTTCCCATCAACGCATACACACGGGACCTTGACACCTTCCCGCGGCCCAACTAT	6575
Db	6361	AGTTTCCCATCAACGCATACACACGGGACCTTGACACCTTCCCGCGGCCCAACTAT	6420
QY	6576	TCCAGGGCGCTATGCGGGTGGCTGTGAGGAGTACGTGGAGGTTACGCGTGTGGGGAT	6635
Db	6421	TCCAGGGCGCTATGCGGGTGGCTGTGAGGAGTACGTGGAGGTTACGCGTGTGGGGAT	6480
QY	6636	TTCCACTACGTGACCGGCGATGACCACTGACAAAGTAAAGTGCCTTCCAGGTTCCGGCC	6695
Db	6481	TTCCACTACGTGACCGGCGATGACCACTGACAAAGTAAAGTGCCTTCCAGGTTCCGGCC	6540
QY	6696	CCCGAATTTTTCACCGAGGTGGATGAGTGGCGTTCACAGGTACGCTCCGGCGTGCAAA	6755
Db	6541	CCCGAATTTTTCACCGAGGTGGATGAGTGGCGTTCACAGGTACGCTCCGGCGTGCAAA	6600
QY	6756	CTCTTCTACGGGAGGACGTCAGTTTCCAGGTCCGGCTCAACCAATACTTGGTTCGGGTGC	6815
Db	6601	CTCTTCTACGGGAGGACGTCAGTTTCCAGGTCCGGCTCAACCAATACTTGGTTCGGGTGC	6660
QY	6816	CAGCTCCCATGCGAGCGGACCGGACGTAACAGTGTCTTACTTCCATGCTCACCGATCCC	6875
Db	6661	CAGCTCCCATGCGAGCGGACCGGACGTAACAGTGTCTTACTTCCATGCTCACCGATCCC	6720
QY	6936	TCCCAATTTACAGCAGAGACGGCTAAGCGTAGGCTGCTAGAGGTTCCCGCTCTTTA	6935
Db	6721	TCCCAATTTACAGCAGAGACGGCTAAGCGTAGGCTGCTAGAGGTTCCCGCTCTTTA	6780
QY	6936	GCCAGTCTACAGTACGACGAGTGTCTGCGCTTCTTTGAAGGCGACATGCACTACCCAC	6995
Db	6781	GCCAGTCTACAGTACGACGAGTGTCTGCGCTTCTTTGAAGGCGACATGCACTACCCAC	6840
QY	6996	CATGACTTCCCGGACGCTGACCTCATCGAGGCCAACCTTGTGGCGGAGGAGATGGGC	7055
Db	6841	CATGACTTCCCGGACGCTGACCTCATCGAGGCCAACCTTGTGGCGGAGGAGATGGGC	6900
QY	7056	GGAAACATCACTCGGTGGAGTCAAGAAATAAGTGTAAATTTTGGACTCTTTTCGAACCG	7115
Db	6901	GGAAACATCACTCGGTGGAGTCAAGAAATAAGTGTAAATTTTGGACTCTTTTCGAACCG	6960
QY	7116	CTTACCGCGAGGGGATGAGAGGAGATATCCGTGCGCGGGAGATCCCTCGGAAAATCC	7175
Db	6961	CTTACCGCGAGGGGATGAGAGGAGATATCCGTGCGCGGGAGATCCCTCGGAAAATCC	7020
QY	7176	AGGAAGTTTCCCTCAGCGTTGCCCATATGGGACGCGCGGACTACAATCTCTCACTGCTA	7235
Db	7021	AGGAAGTTTCCCTCAGCGTTGCCCATATGGGACGCGCGGACTACAATCTCTCACTGCTA	7080
QY	7236	GAGTCTTGGAGGACCCGGACTACGTCCTCCGGTGGTACACGATGCCATGCCACCT	7295
Db	7081	GAGTCTTGGAGGACCCGGACTACGTCCTCCGGTGGTACACGATGCCATGCCACCT	7140
QY	7296	ACCAAGGCTCTTCAATACCACTTCCACGGGAAAGAGGACGGTGTCTGACAGAAATCC	7355
Db	7141	ACCAAGGCTCTTCAATACCACTTCCACGGGAAAGAGGACGGTGTCTGACAGAAATCC	7200
QY	7356	AATGTGTCTTCTGCTGGGAGCTCGCACTAAGACCTTCGCTAGCTCCGGATCGTCG	7415
Db	7201	AATGTGTCTTCTGCTGGGAGCTCGCACTAAGACCTTCGCTAGCTCCGGATCGTCG	7260
QY	7416	GCCGTTGATACGCGACCGGCGACCGCTTCTGCACTGGGCTCCGACGACGGTGACAAA	7475
Db	7261	GCCGTTGATACGCGACCGGCGACCGCTTCTGCACTGGGCTCCGACGACGGTGACAAA	7320
QY	7476	GGATCCGACGTTGAGTGTGCTCTCTCCATGCCCCCTTGAAGGGAGCGCGGGGACCCC	7535

PAI...
IAQ...
LIA...
MO...
AQ...
PV...
AY...
TL...
DS...
MR...
GC...
TD...
AI...
DA...
RS...
DH...
LI...
AL...
OF...
PA...
GL...
IL...
RL...
GW...
AP...
YA...
AR...
VP...
AT...
Y...
DH...
RV...
V...
B...
TC...
DP...
GN...

ORIGIN

Query Match		90.7%; Score 8703.2; DB 13; Length 8780;
Best Local Similarity		99.5%; Pred. No. 0;
Matches 8732; Conservative		0; Mismatches 48; Indels 0; Gaps 0;
Q/	156	CCGGTGAGTACACCGGAATGCCAGGACGACCGGTCCTTTCTTGGATCAACCGGCTCAA 215
D/	1	CCGGTGAGTACACCGGAATGCCAGGACGACCGGTCCTTTCTTGGATCAACCGGCTCAA 60
Q/	216	TGCTCGGAGATTTGGGCGTGCCCCCGGAGACTGCTAGCCGAGTAGTGTGGGTGCGGAA 275
D/	61	TGCTCGGAGATTTGGGCGTGCCCCCGGAGACTGCTAGCCGAGTAGTGTGGGTGCGGAA 120
Q/	276	AGGCTTTGTGTAATGCTGATAGGCTGCTTGGAGTGTGCGAGTGTGCGGAGGCTGCTAGACCG 335
D/	121	AGGCTTTGTGTAATGCTGATAGGCTGCTTGGAGTGTGCGAGTGTGCGGAGGCTGCTAGACCG 180
Q/	336	TGCACCATGAGCAGCAATCTCAAACTCAAGAAAAACCAACGTAACCAACCGCGGC 395
D/	181	TGCACCATGAGCAGCAATCTCAAACTCAAGAAAAACCAACGTAACCAACCGCGGC 240
Q/	396	CCACAGGACGTCAAGTTCCTCGGCGGCTGGTTCAGATCGTTGGTGGAGTTTACCTGTGCGG 455
D/	241	CCACAGGACGTCAAGTTCCTCGGCGGCTGGTTCAGATCGTTGGTGGAGTTTACCTGTGCGG 300
Q/	456	CGCAGGGGCCAGGTTGGGTTGCGCGGACTAGGAAGGCTTCCGAGCGGTGCGAACT 515
D/	301	CGCAGGGGCCAGGTTGGGTTGCGCGGACTAGGAAGGCTTCCGAGCGGTGCGAACT 360
Q/	516	CGTGAAGCGCAGCACTATCCAAAGGCTGCGGACCGGAGGCGAGGCGCTGGGCTCAG 575
D/	361	CGTGAAGCGCAGCACTATCCAAAGGCTGCGGACCGGAGGCGAGGCGCTGGGCTCAG 420
Q/	576	CCCGGGTACCTTTGGCCCTCTATGGCAATGAGGGCTGGGGTGGCGAGGATGGCTCCTG 635

D/	421	CCCGGGTACCTTTGGCCCTCTATGGCAATGAGGGCTTGGGTGGGACGATGGCTCCTG 480
Q/	636	TCAACCCCGGGCTCCCGGCTAGTTGGGCGCCCAACGGAACCCCGGGGTAGGTGCGGTAAAC 695
D/	481	TCAACCCCGGGCTCCCGGCTAGTTGGGCGCCCAACGGAACCCCGGGGTAGGTGCGGTAAAC 540
Q/	696	TTGGGTAAAGTTCATCGATACCTTATCATCGGCTTCCGCGATCTCATGGGTACATTCGG 755
D/	541	TTGGGTAAAGTTCATCGATACCTTATCATCGGCTTCCGCGATCTCATGGGTACATTCGG 600
Q/	756	CTCGTCCGGCGCCCCCTAGGGGCGCTGCGAGGCGCTTGGCACACAGGTGTCCGGTCTG 815
D/	601	CTCGTCCGGCGCCCCCTAGGGGCGCTGCGAGGCGCTTGGCACACAGGTGTCCGGTCTG 660
Q/	816	GAGGACGGGCTGAACATATGCAACAGGGAACCTTGCCTCCGGTGTCTTCTATCTTCTCTC 875
D/	661	GAGGACGGGCTGAACATATGCAACAGGGAACCTTGCCTCCGGTGTCTTCTATCTTCTCTC 720
Q/	876	TTGGCTCTGCTGCTGTTTGGACATCCCGACTTCCGCTTATGAAGTGGCGAACAGTGTCC 935
D/	721	TTGGCTCTGCTGCTGTTTGGACATCCCGACTTCCGCTTATGAAGTGGCGAACAGTGTCC 780
Q/	936	GGGATATACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGAC 995
D/	781	GGGATATACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGAC 840
Q/	996	GTGATCATGATGCTCCCGGTCGCTGTTTTCAGAGGGTAAACAGCTCCCGTTCG 1055
D/	841	GTGATCATGATGCTCCCGGTCGCTGTTTTCAGAGGGTAAACAGCTCCCGTTCG 900
Q/	1056	TGGGTAGCGCTCACTCCAGCGCTCGGGCCAGGAATGCCAGCGTCCCACTACGACAAATA 1115
D/	901	TGGGTAGCGCTCACTCCAGCGCTCGGGCCAGGAATGCCAGCGTCCCACTACGACAAATA 960
Q/	1116	CGACGCCAGCTGACCTTGTCTGTTGGACGCGCTTCTTCTGCTCCGCTATGTAGTGGGG 1175
D/	961	CGACGCCAGCTGACCTTGTCTGTTGGACGCGCTTCTTCTGCTCCGCTATGTAGTGGGG 1020
Q/	1176	GATCTCTGCGGATCTATTTTCTCTCGTCTCCAGCTGTTCACTTCTCGCTCGCGCGGAT 1235
D/	1021	GATCTCTGCGGATCTATTTTCTCTCGTCTCCAGCTGTTCACTTCTCGCTCGCGCGAT 1080
Q/	1236	GAGACGTGAGGACTGCAATCTGCTCAATATATCCCGGCGCATGTATCAGGTACCGCGATG 1295
D/	1081	GAGACGTGAGGACTGCAATCTGCTCAATATATCCCGGCGCATGTATCAGGTACCGCGATG 1140
Q/	1296	GCTTGGGATATGATGAACTGCTACCTACCAACAGCCCTAGTGTGTGCGAGTTGCTC 1355
D/	1141	GCTTGGGATATGATGAACTGCTACCTACCAACAGCCCTAGTGTGTGCGAGTTGCTC 1200
Q/	1356	CGGATCCCAAGCTGCTGGACATGTTGGCGGGGCCCACTGGGGAGTCTTGGCGGC 1415
D/	1201	CGGATCCCAAGCTGCTGGACATGTTGGCGGGGCCCACTGGGGAGTCTTGGCGGC 1260
Q/	1416	CTTGGCTACTATTTCATGTTAGGAACTGGGCTTAAGGTTCTGATTTGGCGCTACTCTTT 1475
D/	1261	CTTGGCTACTATTTCATGTTAGGAACTGGGCTTAAGGTTCTGATTTGGCGCTACTCTTT 1320
Q/	1476	GCGGCGTTGACGGGGAGACCCACACGACGGGAGGTTGGCGGCGCAACACCTCCGGG 1535
D/	1321	GCGGCGTTGACGGGGAGACCCACACGACGGGAGGTTGGCGGCGCAACACCTCCGGG 1380
Q/	1536	TTACAGTCCCTTTTCTCATCTGGGCGCTCTCAGAAAAATCCAGCTTGTGAATACCAACGGC 1595
D/	1381	TTACAGTCCCTTTTCTCATCTGGGCGCTCTCAGAAAAATCCAGCTTGTGAATACCAACGGC 1440
Q/	1596	AGCTGGCACATCAACAGGACTGCCCTAAATTTGCAATGACTCCCTCCAACTGGGTTCTTT 1655
D/	1441	AGCTGGCACATCAACAGGACTGCCCTAAATTTGCAATGACTCCCTCCAACTGGGTTCTTT 1500
Q/	1656	CCCGGCTGTTTATGACACAAAGTTCAACTCGTCCGGTGGCGGAGCGATGGCCAGC 1715

Db 1501 GCCGCGCTGTTTACGCAACAAGTTCAACTCGTCCGGGTGCCCGAGCGCATGCCCAGC 1560
Qy 1716 TCCCGCCCATTTGACTGGTTCCGCCAGAGGGTGGGCCCCCATCACCTATACTAAAGCCTAAC 1775
Db 1561 TCCCGCCCATTTGACTGGTTCCGCCAGAGGGTGGGCCCCCATCACCTATACTAAAGCCTAAC 1620
Qy 1776 AGCTCGGATCAGAGCCCTTATTTGCTGGCAATACGGCCCTCGACCGCTGTGGTGTGTAACC 1835
Db 1621 AGCTCGGATCAGAGCCCTTATTTGCTGGCAATACGGCCCTCGACCGCTGTGGTGTGTAACC 1680
Qy 1836 GGGTGCAGGTGTGGTCCAGTGTATTTGTTTCACCCCAAGCCCTTTGGTGGGGACC 1895
Db 1681 GGGTGCAGGTGTGGTCCAGTGTATTTGTTTCACCCCAAGCCCTTTGGTGGGGACC 1740
Qy 1896 ACCGATCGTTCCGGTGTCCCTACGTATAGCTGGGGGAGAATGAGACAGAGCTGATGCTC 1955
Db 1741 ACCGATCGTTCCGGTGTCCCTACGTATAGCTGGGGGAGAATGAGACAGAGCTGATGCTC 1800
Qy 1956 CTCAACAACAACCGCTCCGCCCAAGGCAACTGGTTCCGCTGTACATGGATGAATAGTACT 2015
Db 1801 CTCAACAACAACCGCTCCGCCCAAGGCAACTGGTTCCGCTGTACATGGATGAATAGTACT 1860
Qy 2016 GGGTTCACTAAGACGTGGAGGTGCCCGTGTAAACATCGGGGGGGTGGTAAACCGCAC 2075
Db 1861 GGGTTCACTAAGACGTGGAGGTGCCCGTGTAAACATCGGGGGGGTGGTAAACCGCAC 1920
Qy 2076 TTGATCTGCCCAACGAGCTGCTTCGGAGAGCACCCGAGGCTACTTACAAAAATGTGGC 2135
Db 1921 TTGATCTGCCCAACGAGCTGCTTCGGAGAGCACCCGAGGCTACTTACAAAAATGTGGC 1980
Qy 2136 TCGGGGCCCTGGTTGACACCTAGTGCCCTAGTAGACTACCCATACAGGCTTTGGCACTAC 2195
Db 1981 TCGGGGCCCTGGTTGACACCTAGTGCCCTAGTAGACTACCCATACAGGCTTCGGCACTAC 2040
Qy 2196 CCTGCACTCTCAATTTTTCATCTTTTAAAGTTAGATGTATGTGGGGGGCGTGGAGCAC 2255
Db 2041 CCTGCACTCTCAATTTTTCATCTTTTAAAGTTAGATGTATGTGGGGGGCGTGGAGCAC 2100
Qy 2256 AGGCTCAATGCGCATGCAATTTGGACTCGAGAGAGCGCTGTAACTTGGAGACAGGGAT 2315
Db 2101 AGGCTCAATGCGCATGCAATTTGGACTCGAGAGAGCGCTGTAACTTGGAGACAGGGAT 2160
Qy 2316 AGGTCAAGAACTCAGCCCGCTGTGCTGTCTCAACAAGAGTGGCAGATATGCCCCGTGCT 2375
Db 2161 AGGTCAAGAACTCAGCCCGCTGTGCTGTCTCAACAAGAGTGGCAGATATGCCCCGTGCT 2220
Qy 2376 TTCACACCCCTACCGGCTTATCCACTGGTTTGAATCCATCTCATCAGAAATCGTGGAC 2435
Db 2221 TTCACACCCCTACCGGCTTATCCACTGGTTTGAATCCATCTCATCAGAACTCGTGGAC 2280
Qy 2436 GTGCAATACCTGTACGGTGTAGGTCAGCGTTGTCTCTCTTGTGCAATCAATGGGAGTAC 2495
Db 2281 GTGCAATACCTGTACGGTGTAGGTCAGCGTTGTCTCTCTTGTGCAATCAATGGGAGTAC 2340
Qy 2496 ATCCTGTGTCTTCTCTCTTGGCAGACGGCGCGTGTGTCCTGTGTTGGATGATG 2555
Db 2341 ATCCTGTGTCTTCTCTCTTGGCAGACGGCGCGTGTGTCCTGTGTTGGATGATG 2400
Qy 2556 CTGCTGATAGCCAGGCTGAGCCGCTTAGAGAACTTGGTGTCTCAATCGGCGTCC 2615
Db 2401 CTGCTGATAGCCAGGCTGAGCCGCTTAGAGAACTTGGTGTCTCAATCGGCGTCC 2460
Qy 2616 GTGGCCGAGCGCATGGTATCTCTCTTCTTGTGTTCTTCTGCGCCGCTGGTACATT 2675
Db 2461 GTGGCCGAGCGCATGGTATCTCTCTTCTTGTGTTCTTCTGCGCCGCTGGTACATT 2520
Qy 2676 AAGGCGAGGCTGCTCTCTGGGCGGCTATGCTTTTTATGGCGTATGGCCGCTGCTCCTG 2735
Db 2521 AAGGCGAGGCTGCTCTCTGGGCGGCTATGCTTTTTATGGCGTATGGCCGCTGCTCCTG 2580
Qy 2736 CTCCTACTGGCGTTACCAACAGAGCTTACGCTTGGACCGGAGATGCTGATCGTGC 2795
Db 2581 CTCCTACTGGCGTTACCAACAGAGCTTACGCTTGGACCGGAGATGCTGATCGTGC 2640

Qy 2796 GGGGTCGGTCTTGTAGTCTGGTATCTTGACTTGTACCATACTACAAAGTGT 2855
Db 2641 GGGGTCGGTCTTGTAGGCTGGTATCTTGACCTTGTACCATACTACAAAGTGT 2700
Qy 2856 CTCACTAGGCTCATATGGTGTACAATACTTTATACACGAGCGGAGCGCACATGCAA 2915
Db 2701 CTCACTAGGCTCATATGGTGTACAATACTTTATACACGAGCGGAGCGCACATGCAA 2760
Qy 2916 GTGTGGGTCCCCCCTCAAGTTTCGGGGAGCGCATGCCATCTCTCTCACGTGT 2975
Db 2761 GTGTGGGTCCCCCCTCAAGTTTCGGGGAGCGCGCATCTCTCTCACGTGT 2820
Qy 2976 GCGGTTTCATCCAGAGTTAAATTTTGTGACATCACCAAACCTCTCTCGCCATCTCGGCCG 3035
Db 2821 GCGGTTTCATCCAGAGTTAAATTTTGTGACATCACCAAACCTCTCTCGCCATCTCGGCCG 2880
Qy 3036 CTATGTGTCTCCAGGCTGGCATAAACGAGAGTGCCTGTCTTCTGGCGCTCAAGGGCTC 3095
Db 2881 CTATGTGTCTCCAGGCTGGCATAAACGAGAGTGCCTGTCTTCTGGCGCTCAAGGGCTC 2940
Qy 3096 ATTCGTGCATGCATGTAGTGCAGAAAGTTCGCGGGGTCAATTATGTCCAAATGTCCTTC 3155
Db 2941 ATTCGTGCATGCATGTAGTGCAGAAAGTTCGCGGGGTCAATTATGTCCAAATGTCCTTC 3000
Qy 3156 ATGAAGCTTGGCGCGCTGACAGGTACGTACGTTTATATACCATCTTACCCCACTCGCGGAC 3215
Db 3001 ATGAAGCTTGGCGCGCTGACAGGTACGTACGTTTATATACCATCTTACCCCACTCGCGGAC 3060
Qy 3216 TGGGCCCAACCGGGCTTACGAGACCTTGGCGTGGCGGTAGAGCCGCTCTTCTTCGCC 3275
Db 3061 TGGGCCCAACCGGGCTTACGAGACCTTGGCGTGGCGGTAGAGCCGCTCTTCTTCGCC 3120
Qy 3276 ATGAGAACCAAGTTCATCCTGGGGAGCAGACCGCTGCTGGGGACATCATCTTG 3335
Db 3121 ATGAGAACCAAGTTCATCCTGGGGAGCAGACCGCTGCTGGGGACATCATCTTG 3180
Qy 3336 GGTCTACCGTCTCCGCCGAAGGGGAAGAGATATTTTGGGACCGGCTGATAGTCTC 3395
Db 3181 GGTCTACCGTCTCCGCCGAAGGGGAAGAGATATTTTGGGACCGGCTGATAGTCTC 3240
Qy 3396 GAAAGGCAAGGTCGGGCTCTTGGCGCCATCAAGGCTTACTCCCAACAAACGCGGGC 3455
Db 3241 GAAAGGCAAGGTCGGGCTCTTGGCGCCATCAAGGCTTACTCCCAACAAACGCGGGC 3300
Qy 3456 GTACTTGGTTCATCATCTAGCTCAGAGCGGGACAGAACCCAGGTGAGGGGAG 3515
Db 3301 GTACTTGGTTCATCATCTAGCTCAGAGCGGGACAGAACCCAGGTGAGGGGAG 3360
Qy 3516 GTTCAAGTGGTTTCTACCGCAACCAATCTTTCTTGGCGACCTGCATCAACGGCGTGTGC 3575
Db 3361 GTTCAAGTGGTTTCTACCGCAACCAATCTTTCTTGGCGACCTGCATCAATGGCGTGTGC 3420
Qy 3576 TGGACTGTCTACCATGGCGCTGCGAAGACCTTAGCCGCTCAAAAGGTCCAATACC 3635
Db 3421 TGGACTGTCTACCATGGCGCTGCGAAGACCTTAGCCGCTCAAAAGGTCCAATACC 3480
Qy 3636 CAATGTACCAATGTAGACTGTGACCTCTGCGGTGGCAGGCGCCCCCGGGCGGCGC 3695
Db 3481 CAATGTACCAATGTAGACTGTGACCTCTGCGGTGGCAGGCGCCCCCGGGCGGCGC 3540
Qy 3696 TCCATGACCAATGTAGACTGTGACCTCTTACTTTTGGTTCACGAGACATCTGAT 3755
Db 3541 TCCATGACCAATGTAGACTGTGACCTCTTACTTTTGGTTCACGAGACATCTGAT 3600
Qy 3756 GTCAATCCGGTTCGCGCGGAGGCGACAGAGGGGAAGTCTACTCTCCCGAGGCCGCTC 3815
Db 3601 GTCAATCCGGTTCGCGCGGAGGCGACAGAGGGGAAGTCTACTCTCCCGAGGCCGCTC 3660
Qy 3816 TCCTACTCTGAAGGCTCTCGGTGGTCCATTTGCTTTTCCCTTGGGACGATCGTGGGC 3875
Db 3661 TCCTACTCTGAAGGCTCTCGGTGGTCCATTTGCTTTTCCCTTGGGACGATCGTGGGC 3720

QY	3876	GTCTTCGGGCTGCTGTGTGCACCCGGGGGGTTCGGAAGCGCGTGAGCTTTCATACCCGTT	3935
Db	3721	GTCTTCGGGCTGCTGTGTGCACCCGGGGGGTTCGGAAGCGCGTGAGCTTTCATACCCGTC	3780
QY	3936	GAGTCTATGGAACCTACCAATGCGGCTCTCCGGTCTTCACAGACAACTCAACCCGCCCGGCT	3995
Db	3781	GAGTCTATGGAACCTACCAATGCGGCTCCCGGCTTTTCAGACAACTCTATCCGCCCGGCT	3840
QY	3996	GTACCGCAGACATTTCAAGTGGCACATCTGCACGCTCTTACTGGCAGCGCAAGACACC	4055
Db	3841	GTACCGCAGACATCCCAGTGGCACATCTGCACGCTCTTACTGGCAGCGCAAGACACT	3900
QY	4056	AAAGTCCCGGCTGGTATGAGCCCAAGGCTACAAGTGTCTGCTGAACCCGCTCGGTT	4115
Db	3901	AAAGTCCCGGCTGGTATGAGCCCAAGGCTACAAGTGTCTGCTGAACCCGCTCGGTT	3960
QY	4116	GCCGCCACCTTAGGGTTGGGGGGTATATGTCCAAGSCACAGTATCGACCCCTAAACATC	4175
Db	3961	GCCGCCACCTTAGGGTTGGGGGGTATATGTCCAAGSCACAGTATCGACCCCTAAACATC	4020
QY	4176	AGAACTGGGGTAAGGACCATTAACACGGGGGGCTTCCATTACGTACTCCACCTATGGCAAG	4235
Db	4021	AGAACTGGGGTAAGGACCATTAACACGGGGGGCTTCCATTACGTACTCCACCTATGGCAAG	4080
QY	4236	TTCTTTGCCAGCGGTGGCTGTTCTGGGGGGCCCTATGACATCATATATGTATGATGTC	4295
Db	4081	TTCTTTGCCAGCGGTGGCTGTTCTGGGGGGCCCTATGACATCATATATGTATGATGTC	4140
QY	4296	CACCTCAACTGACTACCTACCTCTGGGATCGGACAGTCTTGGACCAAGCGGAGACG	4355
Db	4141	CACCTCAACTGACTACCTACCTCTGGGATCGGACAGTCTTGGACCAAGCGGAGACG	4200
QY	4356	GCTGAGCGGGCTGCTGCTGCTCGCCACCGCTACACCTCGGGATCGGTTACCGTGCCA	4415
Db	4201	GCTGAGCGGGCTGCTGCTGCTCGCCACCGCTACACCTCGGGATCGGTTACCGTGCCA	4260
QY	4416	CACCCCAATATCGAGAAATAGGCTGTCCAAATAGGAGAGATCCCTTCTATGGCAAA	4475
Db	4261	CACCCCAATATCGAGAAATAGGCTGTCCAAATAGGAGAGATCCCTTCTATGGCAAA	4320
QY	4476	GCCATCCCATTTAGGCGCATCAAGGGGGGAGGATCTCATTTCTGCCATTCACAAAG	4535
Db	4321	GCCATCCCATTTAGGCGCATCAAGGGGGGAGGATCTCATTTCTGCCATTCACAAAG	4380
QY	4536	AAATGTGACAGCTCGCCGAAAGCTGACAGGCTCGGACCTGGAACGCTGTAGCATATTAC	4595
Db	4381	AAATGTGACAGCTCGCCGAAAGCTGACAGGCTCGGACCTGGAACGCTGTAGCATATTAC	4440
QY	4596	CGGGGCTTGTATGTTCGGTCAATACCGCTATCGGAGACGCTGTTGTGTTGGCAACAGAC	4655
Db	4441	CGGGGCTTGTATGTTCGGTCAATACCGCTATCGGAGAGCTGTTGTGTTGGCAACAGAC	4500
QY	4656	GCTCTAATGACGGGTTTACCGGGGATTTTGAATCAGTGTGAGCTGCAATACATGTGTC	4715
Db	4501	GCTCTAATGACGGGTTTACCGGGGATTTTGAATCAGTGTGAGCTGCAATACATGTGTC	4560
QY	4716	ACCCAGACGTGCACTTCAAGCTTGAATCCACCTTACCATTTGAGACGACACCGTGCC	4775
Db	4561	ACCCAGACGTGCACTTCAAGCTTGAATCCACCTTACCATTTGAGACGACACCGTGCC	4620
QY	4776	CAAGACGGGCTGCTCGGCTCGCAACCGGAGGTAGAACTGGCAGGGGTAGGATGGCATC	4835
Db	4621	CAAGACGGGCTGCTCGGCTCGCAACCGGAGGTAGAACTGGCAGGGGTAGGATGGCATC	4680
QY	4836	TACAGTTTGTGACTCCAGGAAACCGGCCCTGGGCATGTTGCAATCTTCGGTCTGTGT	4895
Db	4681	TACAGTTTGTGACTCCAGGAAACCGGCCCTGGGCATGTTGCAATCTTCGGTCTGTGT	4740
QY	4896	GAGTCTATGACGGGGCTGCTGTTGATGAGCTCAGCCCGCTGAGACCTCGGTTAGG	4955
Db	4741	GAGTCTATGACGGGGCTGCTGTTGATGAGCTCAGCCCGCTGAGACCTCGGTTAGG	4800
QY	4956	TTGCGGGCTTACCTAAATACACAGGGTTGCCGCTGTCGACGACCATCTGAGTTCTTGG	5015

Db	4801	TTGCGGGCTTACCTAAATACACCGGGTTGCCGCTGTCAGACCATCTGAGTTCTGG	4860
QY	5016	GAGAGGCTTTTACAGGCGCTCAACACATAGATGCCCACTTCTGTTCCAGACTAAACAG	5075
Db	4861	GAGAGGCTTTTACAGGCGCTCAACACATAGATGCCCACTTCTGTTCCAGACTAAACAG	4920
QY	5076	GCAGGAGCAACTTTCCTTACCTGTTGGCATATCAAGCTTACAGTGTGCCAGGGCTCAA	5135
Db	4921	GCAGGAGCAACTTTCCTTACCTGTTGGCATATCAAGCTTACAGTGTGCCAGGGCTCAA	4980
QY	5136	GCTTCCACCTCCATCTGTTGGACCAAAATGTGGAAGTGTCTCATACGGCTGAAACCTACACTG	5195
Db	4981	GCTTCCACCTCCATCTGTTGGACCAAAATGTGGAAGTGTCTCATACGGCTGAAACCTACACTG	5040
QY	5196	CACGGGCCAAACCCCTGCTGTATAGGCTAGGAGCGGTCAAAATGAGGTTCATCTCACA	5255
Db	5041	CACGGGCCAAACCCCTGCTGTATAGGCTAGGAGCGGTCAAAATGAGGTTCATCTCACA	5100
QY	5256	CACCCCATTAATTAATACATCATGTCATGTCGCTGACCTGAGGTTCGTCACTAGC	5315
Db	5101	CACCCCATTAATTAATACATCATGTCATGTCGCTGACCTGAGGTTCGTCACTAGC	5160
QY	5316	ACCTGGGTGCTGTAGCGGAGTCTTTCAGAGTTCGCGGCATATCTGCTGACGACGAGC	5375
Db	5161	ACCTGGGTGCTGTAGCGGAGTCTTTCAGAGTTCGCGGCATATCTGCTGACGACGAGC	5220
QY	5376	AGTGTGCTCATTTGTGGGCGAGTCACTTGTTCGGGAAGCGAGCTGTGTTCCCGACAGG	5435
Db	5221	AGTGTGCTCATTTGTGGGCGAGTCACTTGTTCGGGAAGCGAGCTGTGTTCCCGACAGG	5280
QY	5436	GAAGTCTCTTACAGAGTTCGATGAGATGGAAGTGTGCTTCAACAATCTTCTTACATC	5495
Db	5281	GAAGTCTCTTACAGAGTTCGATGAGATGGAAGTGTGCTTCAACAATCTTCTTACATC	5340
QY	5496	GACGAGGAATGAGCTCGCGGAGCAATTCAGCAAAAGCGCTCGGGTGTGTTGCAACG	5555
Db	5341	GACGAGGAATGAGCTCGCGGAGCAATTCAGCAAAAGCGCTCGGGTGTGTTGCAACG	5400
QY	5556	GCCACCAAGCAAGCGGAGCTGCTGCTCCGTTGGAGTCCAAAGTGGCGAGCCCTTGAG	5615
Db	5401	GCCACCAAGCAAGCGGAGCTGCTGCTCCGTTGGAGTCCAAAGTGGCGAGCCCTTGAG	5460
QY	5616	ACCTTCTGCGGAGCAACATGTCGAATTTTATCAGCGGAATACAGTACCTAGCAGGCTTA	5675
Db	5461	ACCTTCTGCGGAGCAACATGTCGAATTTTATCAGCGGAATACAGTACCTAGCAGGCTTA	5520
QY	5676	TCCACTCTGCTCGAAACCCCGGATAGCATCATTTGATGCAATTTACAGCTTCTTACACT	5735
Db	5521	TCCACTCTGCTCGAAACCCCGGATAGCATCATTTGATGCAATTTACAGCTTCTTACACT	5580
QY	5736	AGCCCGCTCACCAACCCCAACCCCTCTGTTTAAACATCTTGGGGGATGGTGGCTGCC	5795
Db	5581	AGCCCGCTCACCAACCCCAACCCCTCTGTTTAAACATCTTGGGGGATGGTGGCTGCC	5640
QY	5796	CAACTCGCTCTCTCCAGCGCTGCTGAGCTTTCGTTGGCGCGGCATCGCGGAGCGGCT	5855
Db	5641	CAGCTCGCTCTCTCCAGCGCTGCTGAGCTTTCGTTGGCGCGGCATCGCGGAGCGGCT	5700
QY	5856	GTTGGGAGCATAGGCTTGGGAAGGTGCTGTCGAGCATCTTGGCGGCTATGGGGCAGGG	5915
Db	5701	GTTGGGAGCATAGGCTTGGGAAGGTGCTGTCGAGCATCTTGGCGGCTATGGGGCAGGG	5760
QY	5916	GTAGCCGCGCATCTGTTGGCTTTTAAAGTTCATGAGCGCGGAGGTGCCCTCCACCGAGAC	5975
Db	5761	GTAGCCGCGCATCTGTTGGCTTTTAAAGTTCATGAGTGGCGAGGTGCCCTCCACCGAGAC	5820
QY	5976	CTGCTCAACTTACTCCCTGCGCATCTCTCTGTTGGTCCCTGCTGCGGGTCTGTGTC	6035
Db	5821	CTGCTCAACTTACTCCCTGCGCATCTCTCTGTTGGTCCCTGCTGCGGGTCTGTGTC	5880
QY	6036	GCAGCAATCTGCTCGGACGTTGGGCGGAGGGGCTGTGCACTGATGAACCGG	6095

Db 5881 GCAGCAATACTCCGTGCGCACGTGGGCCCGGAGAGGAGCTGTGACGTGGATGAACCGG 5940
QY 6096 CTGATAGCGTTCCGTTCCGGGGTAACACAGCTCTCCCTTAGCGACTATGTGCTCAGAGC 6155
Db 5941 CTGATAGCGTTCCGTTCCGGGGTAACACAGCTCTCCCTTAGCGACTATGTGCTCAGAGC 6000
QY 6156 GACGCTGCAGCAGCTGTCACTCAGATCCCTCTAGCCTTACCATCACTCAACTGCTGAAG 6215
Db 6001 GACGCTGCAGCAGCTGTCACTCAGATCCCTCTAGCCTTACCATCACTCAACTGCTGAAG 6060
QY 6216 CCGCTCCACGAGTAATGAGAGCTGCTCTACGCCATGCTCCGGCTCGTGGCTAAGG 6275
Db 6061 CCGCTTCCACGAGTAATGAGAGCTGCTCTAGGCCATGCTCCGGCTCGTGGCTAAGG 6120
QY 6276 GATGTTGGGATTGGATAGCAGGCTGTGACCTTCAAGACCTGCTCCAGTCCAA 6335
Db 6121 GATGTTGGGATTGGATAGCAGGCTGTGACCTTCAAGACCTGCTCCAGTCCAA 6180
QY 6336 CTCTCCCGCGGTACCGGAGTCCCTTTCTGTCTATGCCAACCGGGGTACAAAGGAGTC 6395
Db 6181 CTCTCCCGCGGTACCGGAGTCCCTTTCTGTCTATGCCAACCGGGGTACAAAGGAGTC 6240
QY 6396 TGGCGGGGGACGGCATATGAAACCACTGCCCATGCGGAGCAAGATGCCCGGACAT 6455
Db 6241 TGGCGGGGGACGGCATATGAAACCACTGCCCATGCGGAGCACAGATGCCCGGACAT 6300
QY 6456 GTCAAAACCGTTCCATGAGGATCGTAGGGCTAGAACCTGCAGCAACACGCTGGCACGGA 6515
Db 6301 GTCAAAACCGTTCCATGAGGATCGTAGGGCTAGAACCTGCAGCAACACGCTGGCACGGA 6360
QY 6516 AGCTTCCCATCAACGCATACACCGGACCTTGACACCTCCCGCGGCCCAACTAT 6575
Db 6361 AGCTTCCCATCAACGCATACACCGGACCTTGACACCTCCCGCGGCCCAACTAT 6420
QY 6576 TCAGGGCGCTATGGCGGTGCTGCTGAGAGTACGTGAGGTTACGCGTGTGGGGAT 6635
Db 6421 TCAGGGCGCTATGGCGGTGCTGCTGAGAGTACGTGAGGTTACGCGTGTGGGGAT 6480
QY 6636 TTCCACTAGTACGGGCTGACCACTGACAGGTAAAGTGCCCATGCCAGTTCCGGCC 6695
Db 6481 TTCCACTAGTACGGGCTGACCACTGACAAAGTAAAGTGCCCATGCCAGTTCCGGCC 6540
QY 6696 CCGAATTTCTTACCGAGGTGGATGAGTGGCGTTGCACAGGTAGCTCCGCGCTGCAAA 6755
Db 6541 CCGAATTTCTTACCGAGGTGGATGAGTGGCGTTGCACAGGTAGCTCCGCGCTGCAAA 6600
QY 6756 CTTCTTTACGGAGGACGTCACTGTTCCAGGTGGGCTCAACCAATACTTGTTCGGGTG 6815
Db 6601 CTTCTTTACGGAGGACGTCACTGTTCCAGGTGGGCTCAACCAATACTTGTTCGGGTG 6660
QY 6816 CAGTCCCATGCGNGCCGAACCGGAGTAAAGTGTCTTACTTCCATGCTCACCGATCCC 6875
Db 6661 CAGTCCCATGCGAGCCGAACCGGAGTAAAGTGTCTTACTTCCATGCTCACCGATCCC 6720
QY 6876 TCCCAATTTACAGCAGAGACGGCTAAGCGTAGGCTGGCTAGAGGTTCTCCCGCTCTTTA 6935
Db 6721 TCCCAATTTACAGCAGAGACGGCTAAGCGTAGGCTGGCTAGAGGTTCTCCCGCTCTTTA 6780
QY 6936 GCCAGTCTACGTAGCCAGTTGTCTGCGCTTCTTTGAAGCGGACATGCACTACCCAC 6995
Db 6781 GCCAGTCTACGTAGCCAGTTGTCTGCGCTTCTTTGAAGCGGACATGCACTACCCAC 6840
QY 7055 CATGACTCCCGGACGCTGACCTCATCGAGGCCAACCTTGTGGCGGCGAGGATGGGC 7055
Db 6841 CATGACTCCCGGACGCTGACCTCATCGAGGCCAACCTTGTGGCGGCGAGGATGGGC 6900
QY 7056 GGAACATCACTCCGCTGGAGTCAGAGAATAAGGTAGTAATTTCTGAGCTCTTTTCGAA 7115
Db 6901 GGAACATCACTCCGCTGGAGTCAGAGAATAAGGTAGTAATTTCTGAGCTCTTTTCGAA 6960
QY 7116 CTTTACCGGAGGGGATGAGGGGAGATATCCGTCCGCGCGGAGATCCTGCGAAATCC 7175
Db 6961 CTTTACCGGAGGGGATGAGGGGAGATATCCGTCCGCGCGGAGATCCTGCGAAATCC 7020

QY 7176 AGGAAGTTCCCTCAGCGTTGCCCATATGGGCAAGCCCGGACTACAATCTCTCCACTGCTA 7235
Db 7021 AGGAAGTTCCCTCAGCGTTGCCCATATGGGCAAGCCCGGACTACAATCTCTCCACTGCTA 7080
QY 7236 GAGTCTTGGAAAGAACCCGAGCTAGTCCCTCCGTTGGTGTACAGGATGCCCATTTCCGACCT 7295
Db 7081 GAGTCTTGGAAAGAACCCGAGCTAGTCCCTCCGTTGGTGTGTACAGGATGCCCATTTCCGACCT 7140
QY 7296 ACCAAGCTCTCTCAATACCACTCCACGAGAGGAAAGAGCGGTTGCTTGACAGAAATCC 7355
Db 7141 ACCAAGCTCTCTCAATACCACTCCACGAGAGGAAAGAGCGGTTGCTTGACAGAAATCC 7200
QY 7356 AATGTGTCTTCTGCTTTGGCGGAGCTCGCCACTAAGACCTTTCGGTAGTCTCCGGATCGTCG 7415
Db 7201 AATGTGTCTTCTGCTTTGGCGGAGCTCGCCACTAAGACCTTTCGGTAGTCTCCGGATCGTCG 7260
QY 7416 GCCGTTGATAGCGGACCGGCGCTTCTGCTGACCTTGGCTCCGACGAGCGGTGACAAA 7475
Db 7261 GCCGTTGATAGCGGACCGGCGCTTCTGCTGACCTTGGCTCCGACGAGCGGTGACAAA 7320
QY 7476 GGATCCGAGCTTGAGTCTGCTCTCCATGCCCGCTTGAAGGGGAGCCGGGGACCCC 7535
Db 7321 GGGTCCGAGCTTGAGTCTGCTCTCCATGCCCGCTTGAAGGGGAGCCGGGGACCCC 7380
QY 7536 GATCTCAGCGACGGGTCTTGGTCTACCGTGTAGTGAGGAGCTAGTGAGGATGTCTGTCTGC 7595
Db 7381 GATCTCAGCGACGGGTCTTGGTCTACCGTGTAGTGAGGAGCTAGTGAGGACGTCTGTCTGC 7440
QY 7596 TGCTCAATGTCTATACGTGGACAGGCGCTGTATCATCGCCATCGCTTGCAGGAGAAAT 7655
Db 7441 TGCTCAATGTCTATACGTGGACAGGCGCTGTATCATCGCCATCGCTTGCAGGAGAAAT 7500
QY 7656 AAGCTGCCCATCAACCGCTTGAGCAACTCTTGTGCTGCTGCTCACACACATGCTTACGCC 7715
Db 7501 AAGCTGCCCATCAACCGCTTGAGCAACTCTTGTGCTGCTGCTCACACACATGCTTACGCC 7560
QY 7716 ACAACATCCGCGACGCAAGCTCCGCGAAGAGGTCACTTTTGACAGATTTGCAAGTC 7775
Db 7561 ACAACATCCGCGACGCAAGCTCCGCGAAGAGGTCACTTTTGACAGATTTGCAAGTC 7620
QY 7776 CTGATGATCAATTACCGGACGTACTCAAGGAGATGAAGCGAAGGCTTCCACAGTTAAG 7835
Db 7621 CTGATGATCAATTACCGGACGTACTCAAGGAGATGAAGCGAAGGCTTCCACAGTTAAG 7680
QY 7836 GCTAAGCTTCTATCTATAGAGGAGCTGCAAGCTGACGCGCCCGCACATTTCCGCCAAATCC 7895
Db 7681 GCTAAGCTTCTATCTATAGAGGAGCTGCAAGCTGACGCGCCCGCACATTTCCGCCAAATCC 7740
QY 7896 AAATTTGGCTATGGGCAAGGACGTCGCGAAACCTATCCAGCAGGCGCTTAAACCAATC 7955
Db 7741 AAATTTGGCTATGGGCAAGGACGTCGCGAAACCTATCCAGCAGGCGCTTAAACCAATC 7800
QY 7956 CGCTCCGTGTCGGAGGACTTCTGCTGGAAGACACTGAACACCAATTTGACACCAATCATG 8015
Db 7801 CGCTCCGTGTCGGAGGACTTCTGCTGGAAGACACTGAACACCAATTTGACACCAATCATG 7860
QY 8016 GCAAAAGTGAAGTTTCTGCTGCAACAGAGAGGAGCGCGCAAGCTCGCCTT 8075
Db 7861 GCAAAAGTGAAGTTTCTGCTGCAACAGAGAGGAGCGCGCAAGCTCGCCTT 7920
QY 8076 ATCGTATTTCCAGACCTGGGAGTTGCTGTATGCGAGAAGATGGCCCTTTTACGACGTGGTC 8135
Db 7921 ATCGTATTTCCAGACCTGGGAGTTGCTGTATGCGAGAAGATGGCCCTTTTACGACGTGGTC 7980
QY 8136 TCCACCTTCTCAGCGCGCTGATGGGCTCCTCATACGATTTCAATATCCTCCCGCAAGCAG 8195
Db 7981 TCCACCTTCTCAGCGCGCTGATGGGCTCCTCATACGCGGTTCCAAATACTCCTCCCGCAAGCAG 8040
QY 8196 CCGGTGAGTTCTGCTGTAATACCTGGAAATCAAGAAATGCCCTATGGGCTTCTCATAT 8255
Db 8041 CCGGTGAGTTCTGCTGTAATACCTGGAAATCAAGAAATGCCCTATGGGCTTCTCATAT 8100

1..8780	source	
/organism="Hepatitis C virus"		
/mol_type="genomic RNA"		
/strain="HC-J4"		
/isolate="pcV-J4L9"		
/db_xref="taxon:11103"		
187..>8780	CDS	
/codon_start=1		
/product="polyprotein"		
/protein_id="AAC15731.1"		
/db_xref="GI:3098653"		
/translation="MSTNPKPQRKTKRNTNRRPQDVKPPGGQIVGGVLLPRRQPLR		
GVAPRKASERSQPRGRQPIPKARRPEGRAWAOPGYWPYPLNGELGAGWALLVSPRG		
SRPSWGPTDPRRRSRNLKVIDTLTCGFADLMGXILPYLGA PLGGARALAHGVRLVED		
GVNYATGNDLPKGSFIFLLALLSLCLTIPASATFVRNVSIGIHHVTNDCSNSSIVIEAAD		
VIMHTPGCVPCVREGNSSRCWVALTPILAAARNASVPTTIIIRHVDLLVGTAFACSAHY		
VGDLCSGIFLVLSQLFTEPRRHATVQDCNCSITYPGHVSCHRMAMNDMMNMVPTTALVY		
SHLLRIPOAVVMAGAHGVLVLAAYISMVGNWAKVLIIVALLFAGVDEGTHETTRGVV		
QHTLRISGTFSLSSGASOKIOLVNTNGSMHINRTALNCNDSLQTFGFAALFYAHKFNSS		
GCPEMASCRPIDWFAQWGPPIITYKNSSDORPYCCHYAPRCGVPSVASCQVGPVYQ		
FPFSPVVGTTDRSGAPITYSWGENEITVMLLNTNTRPPQGNWFGPCGTWMSSTGETKCGG		
PPCNTGGVGNRTLICPTDCFRKNEPATYTKCGSGPMLTPRCLVDIPIRLMYHPTCFVNF		
SIYFKVMYVGVGVEHRLNAACTNWRGECNLDRESELSPLLSTSTTEWTLIIPCAFTTL		
PALSTGLTHLHONIVDVOYLVGVSGAFKWEYIILLLFLLADAVACVACLMMWLL		
IAQAAALENIVLNAASVAGAHGILSLVFPFCAAWYIKGRLLPAAVAFYGVWPLLH		
LLALLPVRAYLDRMAAASCGGAVLGLVFLTSPYKYVFLTRLIWMLOQYTRFVPEAR		
MQVWPPPLNVRGGRDAILILITCAVHPELIFDITLKLLAILGLPMLVLOAGITRFVPEVR		
AOGLISACMLVRKVAAGHYIQAMFKLGALTGTYYVNHILTPURDALEGQWRLLAVVE		
PVVFAMETKITWAGDAAACGDIILGLPVSARRGKEIFLPGDASLEGQWRLLASG		
AYSQOTRGVLGVITSLTRGDKNOQVEGVVSTATQSFPLATCINGVCMWTVVHGASG		
TLIAGPKGPITOMYTNVDLQVGAAPPGARSMTPCGSGSSDLYLVTRHADVTPVRRRG		
DSRGLLSRPVSYLKGSGGPIILCPSHGVGVGFRAAVTCRGVAKAVDFIPVESMETTL		
MKSPPVTDNSPFPVQITFOVHALHPGSGSVTKVPAAYAAQGGYKVLVNLNVAARL		
GFGAYMSKAGIDPNIRITGRTIITGGSIITYSGYKFLADGGCGGAYDIIILICDECHS		
TDSTTILIGTGLDQATAGARLVLAATATPGSVTPVPHNTEELGSLNNGBIFFYVK		
ATIPBAIKGGRHLIFCHSKKCDLAALKTGLGLNAYAYRGLDSVPIPIGDVVVGA		
TDALMTGTFDFDSVIDCNTVYTFDPLDPTIETITVQDVAVSRRQRGRG		
RGSIYRFTPTGPRSPGMDPSVLCYCDAGCAWVELTPAETSVRLRAYQNTPLGLPVCO		
DHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAQNTVCARAQAPSPSDQMWKLA		
LIRLKLTHGPTLLYRLGVAQNEVILTHPIIKYIMACMSADLEVSITVSWLVGVGLA		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGALVGVVCAA		
LIRRHVGPGEAVQWNRLIAPASRGNHVSPTVYPESDAAARVLTILSSLTITOLLK		
ALAAQYCLTTSVIVIGRIILSGKAPVDPREVLVEFDEMEBCASVTLVIOEQMLAE		
QFKQKALGLOTATQAAEAAPVSKWRALETFWAKHWNPIISGIYLAGLSTLPGN		
PATISMAFTFASITSLPTHTHTLLNGLGWAAOLAPPASAFAGVAGIAGAAGSI		
GLGKVLVDILAGYAGVAGALVAFKVMGSVSPSTEDLVNLLPAILSPGAL		

Db 61 TGCCTGGAGATTGGGCGTGCCCGCGAGACTGCTAGCCGAGTAGTGTGGGTGCGAA 120
QY 276 AGGCTTGTGTACTGCTGATAGAGTGTCTGCGAGTGCCTCGGAGGTCTCGTAGACG 335
Db 121 AGGCTTGTGTACTGCTGATAGAGTGTCTGCGAGTGCCTCGGAGGTCTCGTAGACG 180
QY 336 TGCACCATGAGCAGCAATCCTAAACCTCAAGAAAAAACAACGTAACACCAACCGCGC 395
Db 181 TGCACCATGAGCAGCAATCCTAAACCTCAAGAAAAAACAACGTAACACCAACCGCGC 240
QY 396 CCACAGGAGCTCAAGTTCCTGGGCGGTGTAGATCTGTGTGGAGTTTACCTGTGCG 455
Db 241 CCACAGGAGCTCAAGTTCCTGGGCGGTGTAGATCTGTGTGGAGTTTACCTGTGCG 300
QY 456 CGCAGGGGCCCCAGGTGGGTGTGCGCGCACTAGGAAGGCTTCGAGGGGTGCGAACT 515
Db 301 CGCAGGGGCCCCAGGTGGGTGTGCGCGCACTAGGAAGGCTTCGAGGGGTGCGAACT 360
QY 516 CGTGAAGCGCAACCTATCCAAAGGCTCGCCGACCCGAGGCGAGGCTTGGGCTCAG 575
Db 361 CGTGAAGCGCAACCTATCCAAAGGCTCGCCGACCCGAGGCGAGGCTTGGGCTCAG 420
QY 576 CCCGGTACCTTTGGCCCCCTATGGCAATGAGGCGCTGGGGTGGGCAAGGATGGCTCCTG 635
Db 421 CCCGGTACCTTTGGCCCCCTATGGCAATGAGGCGCTGGGGTGGGCAAGGATGGCTCCTG 480
QY 636 TCACCCGCGGCTCGGCGCTAGTTGGGCCCCCAGCGACCCCGGCGTAGTTCGGTTAAC 695
Db 481 TCACCCGCGGCTCGGCGCTAGTTGGGCCCCCAGCGACCCCGGCGTAGTTCGGTTAAC 540
QY 696 TTGGGTAAAGTCAATCGATACCTTACATCGCGCTTCGCGCATCTCATGGGGTACATTCG 755
Db 541 TTGGGTAAAGTCAATCGATACCTTACATCGCGCTTCGCGCATCTCATGGGGTACATTCG 600
QY 756 CTCGTGCGGCCCCCTAGGGGCGGTGCGCAGGGCTTGGCACAAGGTTGCGGGTCTG 815
Db 601 CTCGTGCGGCCCCCTAGGGGCGGTGCGCAGGGCTTGGCACAAGGTTGCGGGTCTG 660
QY 816 GAGGACGGGTGAACATATGCAACAGGGAATTTGCGCGGTGTCTTTCTATCTTCCTC 875
Db 661 GAGGACGGGTGAACATATGCAACAGGGAATTTGCGCGGTGTCTTTCTATCTTCCTC 720
QY 876 TTGGCTCTGCTGTCGTGTTGACCATCCAGCTTCGCTTATGAAGTGGCAGCGTCC 935
Db 721 TTGGCTCTGCTGTCGTGTTGACCATCCAGCTTCGCTTATGAAGTGGCAGCGTCC 780
QY 936 GGGATATACCATGTCAACGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 995
Db 781 GGGATATACCATGTCAACGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGGGAC 840
QY 996 GTGATCATGATACTCCGGGTGCGTGCCTGTGTTTCAAGGGGTAAAGCTCCCGTTGC 1055
Db 841 GTGATCATGATACTCCGGGTGCGTGCCTGTGTTTCAAGGGGTAAAGCTCCCGTTGC 900
QY 1056 TGGGTAGCGCTCACTCCAGCTCGCGGCGAGGATGCCAGGTCCTCCACTACGCAATA 1115
Db 901 TGGGTGGCGCTCACTCCAGCTCGCGGCGAGGATGCCAGGTCCTCCACTACGCAATA 960
QY 1116 CGACGCGCAGCTGACCTTGCTGTTGGGACGGGTGCTTTCTGCTCCGCTATGACGTGGG 1175
Db 961 CGACGCGCAGCTGACCTTGCTGTTGGGACGGGTGCTTTCTGCTCCGCTATGACGTGGG 1020
QY 1176 GATCTCTGGGATCTATTTCTGCTGCTCCAGCTGTTCACTTCTCGCTCGCGGAT 1235
Db 1021 GATCTCTGGGATCTATTTCTGCTGCTCCAGCTGTTCACTTCTCGCTCGCGGAT 1080
QY 1236 GAGACAGTGCAGACTGCAACTGCTCAATCTATCCCGGCGATGATCAGGTCAACGCGATG 1295
Db 1081 GCACAGTGCAGACTGCAACTGCTCAATCTATCCCGGCGATGATCAGGTCAACGCGATG 1140
QY 1296 GCTTTGGGATGATGATGAACTGGTCACTTACCAACAGCGCTAGTGTGTCGAGATTGCTC 1355
Db 1141 GCTTTGGGATGATGATGAACTGGTCACTTACCAACAGCGCTAGTGTGTCGCGATTGCTC 1200

QY 1356 CGGATCCCAAGCTGTGTGGACATGTGTGCGGGGCCCCACTTGGGAGTCTTGGCGGGC 1415
Db 1201 CGGATCCCAAGCTGTGTGTGGACATGTGTGCGGGGCCCCACTTGGGAGTCTTGGCGGGC 1260
QY 1416 CTTGCCCTACTATTCCATGTGTAGGAACTGGGCTAAGGTTCTGATGTGTGGCGCTACTCTTT 1475
Db 1261 CTTGCCCTACTATTCCATGTGTAGGAACTGGGCTAAGGTTCTGATGTGTGGCGCTACTCTTT 1320
QY 1476 GCCGGCTGTGACGGGGAGACCCACACAGCGGGGAGGTGGCGGGCCACACACCTCCGGG 1535
Db 1321 GCCGGCTGTGACGGGGAGACCCACACAGCGGGGAGGTGGTCGGCCACACACCTCCGGG 1380
QY 1536 TTCAGCTCCCTTTTCTCATCTGGGGCGTCTCAGAAAAATCCAGCTTGTGAATACCAACGGC 1595
Db 1381 TTCAGTCCCTTTTCTCATCTGGGGCGTCTCAGAAAAATCCAGCTTGTGAATACCAACGGC 1440
QY 1596 AGCTGGCACATCAACAGGACTGCCCTAAATTTGCAATGACTCCCTCCAAACTGGGTCTTT 1655
Db 1441 AGCTGGCACATCAACAGGACTGCCCTAAATTTGCAATGACTCCCTCCAAACTGGGTCTTT 1500
QY 1656 GCCGGCTGTGTTTACGCACACAAAGTTCAACTCGTCCGGTGCCTGGAGCGCATGGCAGC 1715
Db 1501 GCCGGCTGTGTTTACGCACACAAAGTTCAACTCGTCCGGTGCCTGGAGCGCATGGCAGC 1560
QY 1716 TGCCGCCCCATTGACTGTTGCTCGCCAGGGTGGGGCCCCATCACTTACTTAAGCCTAAC 1775
Db 1561 TGCCGCCCCATTGACTGTTGCTCGCCAGGGTGGGGCCCCATCACTTACTTAAGCCTAAC 1620
QY 1776 AGCTCGGATCAGAGGCGCTTATTGCTGGCATTPACGCGCTCGACCGTGTGGTGTGTAACC 1835
Db 1621 AGCTCGGATCAGAGGCGCTTATTGCTGGCATTPACGCGCTCGACCGTGTGGTGTGTAACC 1680
QY 1836 GCGTGCAGAGTGTGTGTCAGTGTATTGTTTACCCCAAGCCCTGTTGTGGTGGGAC 1895
Db 1681 GCGTGCAGAGTGTGTGTCAGTGTATTGTTTACCCCAAGCCCTGTTGTGTGGGAGC 1740
QY 1896 ACCGATCGTTCCGGTGTCCCTACGTTATAGCTGGGGGAGAAATGACAGAGCTGATGCTC 1955
Db 1741 ACCGATCGTTCCGGTGTCCCTACGTTATAGCTGGGGGAGAAATGACAGAGCTGATGCTC 1800
QY 1956 CTCACAAACACGCTCGGCCAACAGGCAACTGGTTCGGCTGTACATGGATGAATAGTACT 2015
Db 1801 CTCACAAACACGCTCGGCCAACAGGCAACTGGTTCGGCTGTACATGGATGAATAGTACT 1860
QY 2016 GGGTTCACTAAGAGCTGCGGAGGTCCCGCTGTAACTATCGGGGGGTGCGTAAACCGAC 2075
Db 1861 GGGTTCACTAAGAGCTGCGGAGGTCCCGCTGTAACTATCGGGGGGTGCGTAAACCGAC 1920
QY 2076 TTGATCTGCCCCACGGACTGTTCCGGAAGCACCCCGAGGCTACTTACACAAAATGTGCG 2135
Db 1921 TTGATCTGCCCCACGGACTGTTCCGGAAGCACCCCTGAGGCTACTTACAAAATGTGCG 1980
QY 2136 TCGGGGCGCTGGTTGACACCTAGGTGCTTAGTACTACCATACAGGCTTTGGCAGTAC 2195
Db 1981 TCGGGGCGCTGGTTGACACCTAGGTGCTTAGTACTACCATACAGGCTTTGGCAGTAC 2040
QY 2196 CCCTGCACTCTCAATTTTTCATCTTTAAGTTTAGATGTATGTGGGGGGGTGAGGAC 2255
Db 2041 CCCTGCACTCTCAATTTTTCATCTTTAAGTTTAGATGTATGTGGGGGGGTGAGGAC 2100
QY 2256 AGGCTCAATGCGGCTATGCAATTTGGACTCGAGGAGCGCTGTACTTTCGAGGACAGGGAT 2315
Db 2101 AGGCTCAATGCGGCTATGCAATTTGGACTCGAGGAGCGCTGTACTTTCGAGGACAGGGAT 2160
QY 2316 AGGTGAGAACTCAGCGCGCTGCTGTCTACAAACAGAGTGGCAGATAGTCCCTGTGCT 2375
Db 2161 AGGTGAGAACTCAGCGCGCTGCTGTCTACAAACAGAGTGGCAGATAGTCCCTGTGCT 2220
QY 2376 TTTACCAACCTACCGGCTTTTATCCATGCTGTTTGTATCCATCTCCATCAGAAACATCGTGAC 2435
Db 2221 TTTACCAACCTACCGGCTTTTATCCATGCTGTTTGTATCCATCTCCATCAGAAACATCGTGAC 2280

QY 2436 GTGCAATACCTGTACGGTGTAGGGTCAGCGTTTGTCTCTCTTTCGAATCAAAATGGGAGTAC 2495
DB 2281 GTGCAATACCTGTACGGTGTAGGGTCAGCGTTTGTCTCTCTTTCGAATCAAAATGGGAGTAC 2340
QY 2496 ATCCTGTGTCTTTCTCTTCTCTGGCAGACGCGCGGTGTGTCCCTGTCTGTCTGTGGATGATG 2555
DB 2341 ATCCTGTGTCTTTCTCTTCTCTGGCAGACGCGCGGTGTGTCTCTGTCTGTGGATGATG 2400
QY 2556 CTGTCTATAGCCAGGCTGAGGCGGCTTAGAGAACTTGGTGTCTCTCAATGGCGGCTCC 2615
DB 2401 CTGTCTATAGCCAGGCTGAGGCGGCTTAGAGAACTTGGTGTCTCTCAATGGCGGCTCC 2460
QY 2616 GTGGCCGAGCGCATGGTATTCTCTCTCTTCTGTGTCTCTCTCTGCGCGCTGTGTACATT 2675
DB 2461 GTGGCCGAGCGCATGGTATTCTCTCTCTTCTGTGTCTCTCTCTGCGCGCTGTGTACATT 2520
QY 2676 AAGGCGAGGCTGCTCTCTGGGCGGGGTATGCTTTTATGGCGTATGGCGGCTGTCTCTG 2735
DB 2521 AAGGCGAGGCTGCTCTCTGGGCGGGGTATGCTTTTATGGCGTATGGCGGCTGTCTCTG 2580
QY 2736 CTCTACTTGGCGTTACCAACAGAGCTTACGCCCTTGGACCGGAGATGGCTGCATCGTGC 2795
DB 2581 CTCTACTTGGCGTTACCAACAGAGCTTACGCCCTTGGACCGGAGATGGCTGCATCGTGC 2640
QY 2796 GGGGGTGGCGTTCTTGTAGTCTGGTATTCTTGTGACCTTGTACCATCTACAAAGTGTIT 2855
DB 2641 GGGGGTGGCGTTCTTGTAGTCTGGTATTCTTGTGACCTTGTACCATCTACAAAGTGTIT 2700
QY 2856 CTCACTAGGCTCATATGGTGGTTACATATCTTTATCACAGACCGGAGGCGCATGSCAA 2915
DB 2701 CTCACTAGGCTCATATGGTGGTTACATATCTTTATCACAGACCGGAGGCGCATGSCAA 2760
QY 2916 GTGTGGTCCGCCCTTCAACGTTTGGGGAGGCGCGATGCCATCATCTCTCTCACGTGT 2975
DB 2761 GTGTGGTCCGCCCTTCAACGTTTGGGGAGGCGCGATGCCATCATCTCTCTCACGTGT 2820
QY 2976 GCGGTTTCATCAGAGTTAAATTTTGTGACATCACCAAACTCTGTCTGCGCATACTCGGCCCG 3035
DB 2821 GCGGTTTCATCAGAGTTAAATTTTGTGACATCACCAAACTCTGTCTGCGCATACTCGGCCCG 2880
QY 3036 CTATGGTGTCCAGGCTGCATTAACGAGATGCGGTACTTCTGCGCGCTCAAGGCGTC 3095
DB 2881 CTATGGTGTCCAGGCTGCATTAACGAGATGCGGTACTTCTGCGCGCTCAAGGCGTC 2940
QY 3096 ATTCTGTGCATGTAGTTCGAAAAAGTGGCGGGGTCAATATGTCCAAATGGTCTTC 3155
DB 2941 ATTCTGTGCATGTAGTTCGCGAAAGTGGCGGGGTCAATATGTCCAAATGGCGCTTC 3000
QY 3156 ATGAAGCTGGGCGCGCTGACAGGTACGTACGTTTATTAACCAATCTTACCCCATCTGCGGGAC 3215
DB 3001 ATGAAGCTGGGCGCGCTGACAGGTACGTACGTTTATTAACCAATCTTACCCCATCTGCGGGAC 3060
QY 3216 TGGGCCACCGGGCGCTTACGAGACTTTCGGGTGGCGGTAGAGCCGCTCTCTCCGCC 3275
DB 3061 TGGGCCACCGGGCGCTTACGAGACTTTCGGGTGGCGGTAGAGCCGCTCTCTCCGCC 3120
QY 3276 ATGAGAGCAAGGTTCATCACCTGGGAGACGACACCGCTGCGTGTGGGACATCATCTTG 3335
DB 3121 ATGAGAGCAAGGTTCATCACCTGGGAGACGACACCGCTGCGTGTGGGACATCATCTTG 3180
QY 3336 GGCTACCCGCTCTCCGCCGAAGGGGAGAGATATTTTGGACCGGCTGTAGTCTC 3395
DB 3181 GGCTACCCGCTCTCCGCCGAAGGGGAGAGATATTTTGGGACCGGCTGTAGTCTC 3240
QY 3396 GAAGGGCAAGGTTGGCGACTCTCTTGGCCCATACAGGCTTACTCCCAACAACCGGGGC 3455
DB 3241 GAAGGGCAAGGTTGGCGACTCTCTTGGCCCATACAGGCTTACTCCCAACAACCGGGGC 3300
QY 3456 GTACTTGGTGCATCATCTAGCTCACAGGCGGGACAGAACCGAGTTCGAAGGGGAG 3515
DB 3301 GTACTTGGTGCATCATCTAGCTCACAGGCGGGACAGAACCGAGTTCGAAGGGGAG 3360
QY 3516 GTTCAAGTGGTTTCTTACCGCAACAAATCTTTTCTTGGCGACCTTGCAATCAACGGGCTGTGC 3575

DB 3361 GTTCAAGTGGTTTCTTACCGCAACAAATCTTTCTTGGCGACCTTGCAATTAACGCGTGTGC 3420
QY 3576 TGGACTGTCTACATGGCGCTGGCTCGAAGACCTTAGCGCGTCCAAAAGGTCCAATCACC 3635
DB 3421 TGGACTGTCTACATGGCGCTGGCTCGAAGACCTTAGCGCGTCCAAAAGGTCCAATCACC 3480
QY 3636 CAAAATGACCAATGTAGACCTTGGACCTTCTGCGGTGCGAGCGCGCCCGGGGGCGCG 3695
DB 3481 CAAAATGACCAATGTAGACCTTGGACCTTCTGCGGTGCGAGCGCGCCCGGGGGCGCG 3540
QY 3696 TCCATGACACCAATGTAGCTGTGGCAGCTCGGACCTTTACTTTGGTTCAGAGACATGCTGAT 3755
DB 3541 TCCATGACACCAATGTAGCTGTGGCAGCTCGGACCTTTACTTTGGTTCAGAGACATGCTGAT 3600
QY 3756 GTCAATTCGCGTGGCGGAGGCGAGCAGAGGGGAAAGTCTACTCTCCCGCAGGCGCGTC 3815
DB 3601 GTCAATTCGCGTGGCGGAGGCGAGCAGCAGGGGAAAGTCTACTCTCCCGCAGGCGCGTC 3660
QY 3816 TCCATCTCAAGGCTCTCTGGGTGGTCCATTTGCTTTGCGCTTCGGGGCAGCTGTGGGC 3875
DB 3661 TCCATCTCAAGGCTCTCTGGGTGGTCCATTTGCTTTTGGCCCTTCGGGGCAGCTGTGGGC 3720
QY 3876 GTCTTCCGGCTGCTGTGTGCACCCCGGGGGTCCGAAAGCGGTGGACTTTCATACCCGTT 3935
DB 3721 GTCTTCCGGCTGCTGTGTGCACCCCGGGGGTCCGAAAGCGGTGGACTTTCATACCCGTT 3780
QY 3936 GAGTCTATGGAACTACCATGTGGTCTCCGGTCTTCAAGACAACTCAACCCCCCGGCT 3995
DB 3781 GAGTCTATGGAACTACCATGTGGTCTCCCGTCTTCAAGACAACTCATCCCCCGGCT 3840
QY 3996 GTACCGGAGACATTCGAAGTGGCAGCTTGCAGCGCTCTTACGCGAGCGGCAAGAGACC 4055
DB 3841 GTACCGGAGACATTCGAAGTGGCAGCTTGCAGCGCTCTTACGCGAGCGGCAAGAGACT 3900
QY 4056 AAAGTCCGCGTGTGAGCGCCCAAGGTGCTCAAGGTGCTCGTCTGAAACCGCTCCGTT 4115
DB 3901 AAAGTCCGCGTGTGAGCGCCCAAGGTGCTCAAGGTGCTCGTCTGAAACCGCTCCGTT 3960
QY 4116 GCGGCCACCTTAGGGTTTGGGGCGTATATGTCCAAGGCAACCGGTATCGACCTTAACATC 4175
DB 3961 GCGGCCACCTTAGGGTTTGGGGCGTATATGTCCAAGGCAACCGGTATCGACCTTAACATC 4020
QY 4176 AGAATGGGGTAAGACCAATTCACGCGCGGCTCCATACGTAAGTCTTACCTTACGCGCAAG 4235
DB 4021 AGAATGGGGTAAGACCAATTCACGCGCGGCTCCATACGTAAGTCTTACCTTACGCGCAAG 4080
QY 4236 TTCTTTGCCGACGCTGCTGTCTGGGGCGCTTATGACATCATATAATGTGATGAGTGC 4295
DB 4081 TTCTTTGCCGACGCTGCTGTCTGGGGCGCTTATGACATCATATAATGTGATGAGTGC 4140
QY 4296 CACTCAACTGACTCGACTACCATCTTTGGGCGATCGGCAAGTCTTGGACCAAGCGGAGACG 4355
DB 4141 CATTCAACTGACTCGACTACCATCTTTGGGCGATCGGCAAGTCTTGGACCAAGCGGAGACG 4200
QY 4356 GCTGGAGCGCGCTGTGTGTCTGCGCACCGCTACACCTCGGGATCGGTACCGTGCCA 4415
DB 4201 GCTGGAGCGCGCTGTGTGTCTGCGCACCGCTACACCTCGGGATCGGTACCGTGCCA 4260
QY 4416 CACCCCAATTCGAGGAAATAGGCTGTCCAAATGGAGAGATCCCTCTTATGCGAAA 4475
DB 4261 CACCCCAATTCGAGGAAATAGGCTGTCCAAATGGAGAGATCCCTCTTATGCGAAA 4320
QY 4476 GCATCCCCATTTAGGGCGATCAAGGGGGGAGGATCTCTCAATTTTCTGCCATTTCAAGAAG 4535
DB 4321 GCATCCCCATTTAGAGCCATCAAGGGGGGAGGATCTCTCAATTTTCTGCCATTTCAAGAAG 4380
QY 4536 AAATGTGACGAGCTCGCGCAAAAGCTGACGGCTCGGAAGTGAACGCTGTAGCATATTAC 4595
DB 4381 AAATGTGACGAGCTCGCGCAAAAGCTGACGGCTCGGAAGTGAACGCTGTAGCATATTAC 4440
QY 4596 CGGGGCTTGTATGTCTCGTCAATACCGCTTACGAGACGCTGTGTCTGCGGCAACAGAC 4655

Db 4441 CGGGCCCTTGATGTGTCGCTCATACCCGCTATCGGAGAGCGTGTGTGTCGGCAACAGAC 4500
QY 4656 GCTCTAATGACGGGTTTACCGGCCATTTTGACTCAGTGATCGACTGCAATACATATGTGTC 4715
Db 4501 GCTCTAATGACGGGTTTACCGGCCATTTTGACTCAGTGATCGACTGCAATACATATGTGTC 4560
QY 4716 ACCCAGACAGTCGACTTCAGCTTTGATCCCACTTTCACCATTTGAGACGACGACGTCGCC 4775
Db 4561 ACCCAGACAGTCGACTTCAGCTTTGATCCCACTTTCACCATTTGAGACGACGACGTCGCC 4620
QY 4776 CAAAGACGGGTGTCGCGCTCGCAACCGGAGAGTAGAACTAGGACGGGTGAGAGTGGCAATC 4835
Db 4621 CAAAGACGGGTGTCGCGCTCGCAACCGGAGAGTAGAACTAGGACGGGTGAGAGTGGCAATC 4680
QY 4836 TACAGGTTTGTGACTCCAGGAGAA CGGCCCTCGGSCATGTTTCGATTTCTTCGTCCTGTGT 4895
Db 4681 TACAGGTTTGTGACTCCAGGAGAA CGGCCCTCGGSCATGTTTCGATTTCTTCGTCCTGTGT 4740
QY 4896 GAGTGTATGACGCGGGCTGTGCTTTGGTATGAGCTCACGCCGCTGAGACCTCGGTTAGG 4955
Db 4741 GAGTGTATGACGCGGGCTGTGCTTTGGTATGAGCTCACGCCGCTGAGACCTCGGTTAGG 4800
QY 4956 TTGCGGGCTTACTAAATACACAGGGTTGCCGTCTGCGAGGACATCTGAGATTCTGG 5015
Db 4801 TTGCGGGCTTACTAAATACACAGGGTTGCCGTCTGCGAGGACATCTGAGATTCTGG 4860
QY 5016 GAGAGGCTTTCACAGGCTCACACATAGATGCCACTTCCCTGCCAGACTAAACAG 5075
Db 4861 GAGAGGCTTTCACAGGCTCACACATAGATGCCACTTCCCTGCCAGACTAAACAG 4920
QY 5076 GCAGGAGACAACTTTCCTTACCTGTGGCATATCAAGCTACAGTGTGGCCAGGGCTCAA 5135
Db 4921 GCAGGAGACAACTTTCCTTACCTGTGGCATATCAAGCTACAGTGTGGCCAGGGCTCAA 4980
QY 5136 GTCCACCTCCATCTGTGGGACAAATGTGGAAGTGTCTCATACGCTGAAACCTACATG 5195
Db 4981 GCTCCACCTCCATCTGTGGGACAAATGTGGAAGTGTCTCATACGCTGAAACCTACATG 5040
QY 5196 CAGGGCCCAACCCCTGCTGTATAGGCTAGAGAGCGGTCCAAATGAGTCAATCTCACA 5255
Db 5041 CAGGGCCCAACCCCTGCTGTATAGGCTAGAGAGCGGTCCAAATGAGTCAATCTCACA 5100
QY 5256 CACCCCATAACTAAATACATATGATGCAATGTGGCTGACCTGGAGTCTGCTACTAGC 5315
Db 5101 CACCCCATAACTAAATACATATGATGCAATGTGGCTGACCTGGAGTCTGCTACTAGC 5160
QY 5316 ACTGGGTGCTGGTAGGCGAGTCTCTTGAGCTTTGGCGCATATGCGCTGACGACAGC 5375
Db 5161 ACTGGGTGCTGGTAGGCGAGTCTCTTGAGCTTTGGCGCATATGCGCTGACGACAGC 5220
QY 5376 AGTGTGTCATTTGCGGAGGATCATCTTGTCCGGGAGCGAGCTGTGTTCCCGACAGG 5435
Db 5221 AGTGTGTCATTTGCGGAGGATCATCTTGTCCGGGAGCGAGCTGTGTTCCCGACAGG 5280
QY 5436 GAAGTCTCTACAGAGTTCGATGAGATGGAAGAGTGTGCTCACAACCTTCTTACATC 5495
Db 5281 GAAGTCTCTACAGAGTTCGATGAGATGGAAGAGTGTGCTCACAACCTTCTTACATC 5340
QY 5496 GAGCAGGGAATGACGCTCCCGAGCAATTTCAAGCAAAAGCGCTCGGGTTGTGCAAAAG 5555
Db 5341 GAGCAGGGAATGACGCTCCCGAGCAATTTCAAGCAAAAGCGCTCGGGTTGTGCAAAAG 5400
QY 5556 GCCACCAAGCAAGCGAGGCTGCTGCTCCGCTGGTGGAGTCCAGTGGCGAGCCCTTGAG 5615
Db 5401 GCCACCAAGCAAGCGAGGCTGCTGCTCCGCTGGTGGAGTCCAGTGGCGAGCCCTTGAG 5460
QY 5616 ACTTCTGCGGCAAGCACATGTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGCTTA 5675
Db 5461 ACTTCTGCGGCAAGCACATGTGGAATTTTCATCAGCGGAATACAGTACCTAGCAGCTTA 5520
QY 5676 TCCACTCTGCTGGAAACCCCGATAGCATCATTTGATGGCAATTTACAGTCTTATCACT 5735
Db 5521 TCCACTCTGCTGGAAACCCCGATAGCATCATTTGATGGCTTTTACAGCTTCTATCACT 5580

QY 5736 AGCCGCTCACCAACCCCAACACCCCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCC 5795
Db 5581 AGCCGCTCACCAACCCCAACACCCCTCTCTGTTTAAACATCTTGGGGGATGGGTGGCTGCC 5640
QY 5796 CAACCTGCTCTCCAGCGTGTGCTAGCTTTCTGTGGGCGCGCATCGCCGAGACGGCT 5855
Db 5641 CAACCTGCTCTCCAGCGTGTGCTAGCTTTCTGTGGGCGCGCATCGCCGAGACGGCT 5700
QY 5856 GTTGGCAGCATAGGCTTTGGGAGGTCGTCTGAGACATCTTGGCGGCTATGGGCGAGG 5915
Db 5701 GTTGGCAGCATAGGCTTTGGGAGGTCGTCTGAGACATCTTGGCGGCTATGGGCGAGG 5760
QY 5916 GTAGCCCGCGCACTCTGTGGCTTTTAAAGTTCATGAGCGCGAGGTGCCCTCCACCGAGAC 5975
Db 5761 GTAGCCCGCGCACTCTGTGGCTTTTAAAGTTCATGAGCGCGAGGTGCCCTCCACCGAGAC 5820
QY 5976 CTGTCAACTTATCTCCCTGCGCATCTCTCTCTGCTGGTCCCTGCTGCGGGTCTGTGTC 6035
Db 5821 CTGTCAACTTATCTCCCTGCGCATCTCTCTCTGCTGGTCCCTGCTGCGGGTCTGTGTC 5880
QY 6036 GCAGCAATCTGCTCGGCACTGGGCGCGGAGAGGGGCTGTGCACTATGTCCCTGAGAGC 6095
Db 5881 GCAGCAATCTGCTCGGCACTGGGCGCGGAGAGGGGCTGTGCACTATGTCCCTGAGAGC 5940
QY 6096 CTGATAGCTTCTGCTCGGGGTAAACACGCTCTCCCTACGCACTATGTCCCTGAGAGC 6155
Db 5941 CTGATAGCTTCTGCTCGGGGTAAACACGCTCTCCCTACGCACTATGTCCCTGAGAGC 6000
QY 6156 GACGCTGACGACGCTGTCACTCAGATCTCTCTAGCTTTACCATCACTCAACTGCTGAAG 6215
Db 6001 GACGCTGACGACGCTGTCACTCAGATCTCTCTAGCTTTACCATCACTCAACTGCTGAAG 6060
QY 6216 CCGCTCACAGTGGATTAATGAGGACTGCTCTAGCCCATGCTCCGGCTCGTGGCTAAG 6275
Db 6061 CCGCTTACCAGTGGATTAATGAGGACTGCTCTAGCCCATGCTCCGGCTCGTGGCTAAG 6120
QY 6276 GATGTTGGGATTCGATATGACGCTGTGACTGACTTCAAGACTGTGCTCCAGTCCAAA 6335
Db 6121 GATGTTGGGATTCGATATGACGCTGTGACTGACTTCAAGACTGTGCTCCAGTCCAA 6180
QY 6336 CTCTGCGCGGTTTACCGGAGTCCCTTTCTGTCTATGCCAACCGGGTACAAGGAGTC 6395
Db 6181 CTCTGCGCGGTTTACCGGAGTCCCTTTCTGTCTATGCCAACCGGGTACAAGGAGTC 6240
QY 6396 TGGCGGGGAGCGGATCATGCCAACCTCTGCCCATGCGGAGCAGATCGCCGAGAT 6455
Db 6241 TGGCGGGGAGCGGATCATGCCAACCTCTGCCCATGCGGAGCAGATCGCCGAGAT 6300
QY 6456 GTCAAAAACGGTTCATGAGGATCGTAGGGCTTAGAACCTCGAGCAACACGTTGGCACGGA 6515
Db 6301 GTCAAAAACGGTTCATGAGGATCGTAGGGCTTAGAACCTCGAGCAACACGTTGGCACGGA 6360
QY 6516 AGTTCCTCCATCAACGCGATACCAACCGGACTCTTGACACCTTCCCGGGGCCAACTAT 6575
Db 6361 AGTTCCTCCATCAACGCGATACCAACCGGACTCTTGACACCTTCCCGGGGCCAACTAT 6420
QY 6576 TCCAGGGCGCTATGGCGGGTGGCTGCTGAGAGTACGTTGAGAGTTCAGCGGTGCGGGAT 6635
Db 6421 TCCAGGGCGCTATGGCGGGTGGCTGCTGAGAGTTCAGCGGTGCGGGAT 6480
QY 6636 TTCCACTACGTGACGGGATGACCACTGACAACTGAAAGTCCCATGCGAGTTCGGGCC 6695
Db 6481 TTCCACTACGTGACGGGATGACCACTGACAACTGAAAGTCCCATGCGAGTTCGGGCC 6540
QY 6696 CCCGAATTTCTTACGGAGGTGGAGTGGGTTGCAAGGTACGCTCCGCGCTGCAAA 6755
Db 6541 CCCGAATTTCTTACGGAGGTGGAGTGGGTTGCAAGGTACGCTCCGCGCTGCAAA 6600
QY 6756 CTTCTTACGGGAGGAGTCAAGTTCAGGTTCGGGCTCAACCAATATCTTGTTCGGTTCG 6815
Db 6601 CTTCTTACGGGAGGAGTCAAGTTCAGGTTCGGGCTCAACCAATATCTTGTTCGGTTCG 6660

QY	6816	CAGCTCCCATGCGAGCCCGAACC	CGGACGTAA	CAGTGCCTTACTTCCATGCTCACC	GATCCC	6875
Db	6661	CAGCTCCCATGCGAGCCCGAACC	CGGACGTAA	CAGTGCCTTACTTCCATGCTCACC	GATCCC	6720
QY	6876	TCCACATTTACAGCAGACGGCTA	AGCGTAGGCTAGAGGGTCTCC	CCCTCTTTA		6935
Db	6721	TCCACATTTACAGCAGACGGCTA	AGCGTAGGCTAGAGGGTCTCC	CCCTCTTTA		6780
QY	6936	GCCAGCTCATAGCTAGCCAGTTGT	CTGCGCTTCTTTGAAGCGACATGCA	TACCCAC		6995
Db	6781	GCCAGCTCATAGCTAGCCAGTTGT	CTGCGCTTCTTTGAAGCGACATGCA	TACCCAC		6840
QY	6996	CATGACTCCCGGACCGTCACTCAT	CGAGCCAACTCTTGGCGGCGAGGATGGC			7055
Db	6841	CATGACTCCCGGACCGTCACTCAT	CGAGCCAACTCTTGGCGGCGAGGATGGC			6900
QY	7056	GGAACATCACTCGCTGGAGTCAG	AGATAAAGTAGTAATCTGCACTCTT	TCGAACCG		7115
Db	6901	GGAACATCACTCGCTGGAGTCAG	AGATAAAGTAGTAATCTGCACTCTT	TCGAACCG		6960
QY	7116	CTTCACGCGAGGGGATGAGAGG	GAGATATCCGTCGCGCGAGATCCT	TGCGAAATCC		7175
Db	6961	CTTCACGCGAGGGGATGAGAGG	GAGATATCCGTCGCGCGAGATCCT	TGCGAAATCC		7020
QY	7176	AGAAAGTTCCTCAGCGTTGCCAT	ATAGGGACGCGCGGACTACAA	TCTCCACTGCTA		7235
Db	7021	AGAAAGTTCCTCAGCGTTGCCAT	ATAGGGACGCGCGGACTACAA	TCTCCACTGCTA		7080
QY	7236	GAGTCTGGAAGACCGGACTACG	TCCCTCCGTTGTACAGGATGCG	CACTGCCACCT		7295
Db	7081	GAGTCTGGAAGACCGGACTACG	TCCCTCCGTTGTACAGGATGCG	CACTGCCACCT		7140
QY	7296	ACCAAGGCTCCTCCAAATACCA	CTCCACGAGAAAGAGGACGGT	TGCTTGACAGATCC		7355
Db	7141	ACCAAGGCTCCTCCAAATACCA	CTCCACGAGAAAGAGGACGGT	TGCTTGACAGATCC		7200
QY	7356	AATGTGTCTTCTGCGGAGCTCG	CCACTAAGACCTTCGGTAGCT	CCGATCGTCG		7415
Db	7201	AATGTGTCTTCTGCGGAGCTCG	CCACTAAGACCTTCGGTAGCT	CCGATCGTCG		7260
QY	7416	GCGTTGATAGCGGACGGCGCT	CTTCCGCTGACCTCCGACGCG	TGACAA		7475
Db	7261	GCGTTGATAGCGGACGGCGCT	CTTCCGCTGACCTCCGACGCG	TGACAA		7320
QY	7476	GGATCCGACGTTGAGTCGTACT	CTCTCCATGCCCCCTTGAAG	GGGAGCCGGGGACCCC		7535
Db	7321	GGATCCGACGTTGAGTCGTACT	CTCTCCATGCCCCCTTGAAG	GGGAGCCGGGGACCCC		7380
QY	7536	GATCTCAGCGACGGGCTTGGT	CTACCGTAGTGAGAGGCTAG	TGAGGATGCTGCTGC		7595
Db	7381	GATCTCAGCGACGGGCTTGGT	CTACCGTAGTGAGAGGCTAG	TGAGGATGCTGCTGC		7440
QY	7596	TGCTCAATGCTTATAGTGAC	AGGCGCCCTGATCAGCCATG	CGCTGCGGAGGAAGT		7655
Db	7441	TGCTCAATGCTTATAGTGAC	AGGCGCCCTGATCAGCCATG	CGCTGCGGAGGAAGT		7500
QY	7656	AAGTCCGCTCAACCGTTGAG	CAACTCTTTGCTGCTCACC	CAACATGGTCTACGCC		7715
Db	7501	AAGTCCGCTCAACCGTTGAG	CAACTCTTTGCTGCTCACC	CAACATGGTCTACGCC		7560
QY	7716	ACAAATCCCGCAGCGCAAG	CTCCGCGAGAAGGTCACCT	TTTGAAGTTCGAAGTC		7775
Db	7561	ACAAATCCCGCAGCGCAAG	CTCCGCGAGAAGGTCACCT	TTTGAAGTTCGAAGTC		7620
QY	7776	CTGATGATCATTAACCGG	AGCTATCAAGGAGATGAAG	GGGAGGCGTCCACAGTTAAG		7835
Db	7621	CTGATGATCATTAACCGG	AGCTATCAAGGAGATGAAG	GGGAGGCGTCCACAGTTAAG		7680
QY	7836	GCTAAGCTTCTATATAG	AGGCGCTGCAAGCTGAC	CGCCCCACATTCGGGCCAAATCC		7895
Db	7681	GCTAAGCTTCTATATAG	AGGCGCTGCAAGCTGAC	CGCCCCACATTCGGGCCAAATCC		7740
QY	7896	AAATTTGGCTATGGGCA	AAAGGAGCTCCGGAACTAT	TCCAGCAGGGCGGTTAAC	CACATC	7955

RESULT 13

Db	7741	AAATTTGGCTATGGGCA	AAAGGAGCTCCGAACTAT	TCCAGCAGGGCCGTTAAC	CACATC	7800
QY	7956	CGCTCGTGTGGAGGACTT	GCTGGAAGACACTGAA	ACCAATTGACACCA	CCATCATG	8015
Db	7801	CGCTCGTGTGGAGGACTT	GCTGGAAGACACTGAA	ACCAATTGACACCA	CCATCATG	7860
QY	8016	GCAAAAGTGAGGTTTCT	CGCTCCAAACAGAGAGG	AGCGCAAGCCAGCT	TCGCTT	8075
Db	7861	GCAAAAGTGAGGTTTCT	CGCTCCAAACAGAGAGG	AGCGCAAGCCAGCT	TCGCTT	7920
QY	8076	ATCGTATTTCCAGAC	CTCGGAGTTCTGATG	CGAGAAGATGCC	CTTTACGACGTGGTC	8135
Db	7921	ATCGTATTTCCAGAC	CTCGGAGTTCTGATG	CGAGAAGATGCC	CTTTACGACGTGGTC	7980
QY	8136	TCACCCCTTCTCAG	CCCTGATGGGCTCT	CTATACGATTTCA	TACTCCCCCAAGCAG	8195
Db	7981	TCACCCCTTCTCAG	CCCTGATGGGCTCT	CTATACGATTTCA	TACTCCCCCAAGCAG	8040
QY	8196	CGGTCGAGTTCTCT	GGTGAATACCTGG	GAATCAAGAAATGCC	CTATGGCTTCTCATAT	8255
Db	8041	CGGTCGAGTTCTCT	GGTGAATACCTGG	GAATCAAGAAATGCC	CTATGGCTTCTCATAT	8100
QY	8256	GACACCGCTGTTT	TGACTCAACGGT	CACTGAGAGTGAC	ATTCTGTTGAGGAGTCAAT	8315
Db	8101	GACACCGCTGTTT	TGACTCAACGGT	CACTGAGAGTGAC	ATTCTGTTGAGGAGTCAAT	8160
QY	8316	TACCAATGTTG	ATGACTTGGCCCG	AGGCGACAGCC	ATAAGTTCGCTC	8375
Db	8161	TACCAATGTTG	ATGACTTGGCCCG	AGGCGACAGCC	ATAAGTTCGCTC	8220
QY	8376	CTTTACATCGGG	GTCCCTGACTACT	CAAAAGGCGA	NACTCGGTTATCCCGGTGC	8435
Db	8221	CTTTACATCGGG	GTCCCTGACTACT	CAAAAGGCGA	NACTCGGTTATCCCGGTGC	8280
QY	8436	CGGCAAGTGG	CTGTGACGACTAG	CTGCGGTAAT	ACCTCACATGTTTACTTTGAAGGCC	8495
Db	8281	CGGCAAGTGG	CTGTGACGACTAG	CTGCGGTAAT	ACCTCACATGTTTACTTTGAAGGCC	8340
QY	8496	ACTGACGCTGT	CGAGCTGCAAG	AGCTCCAGGACTG	CAACGATGCTG	8555
Db	8341	ACTGACGCTGT	CGAGCTGCAAG	AGCTCCAGGACTG	CAACGATGCTG	8400
QY	8556	CTTGTGCTTAT	CTGTGAAGCGGG	GAACCCAGGAGAT	CGGGGCCCTACGACGCTTC	8615
Db	8401	CTTGTGCTTAT	CTGTGAAGCGGG	GAACCCAGGAGAT	CGGGGCCCTACGACGCTTC	8460
QY	8616	ACGAGGCTAT	GACTAGGTAT	TCGCCCCCGGGGAT	CCGCCCAACAGATA	8675
Db	8461	ACGAGGCTAT	GACTAGGTAT	TCGCCCCCGGGGAT	CCGCCCAACAGATA	8520
QY	8676	CTGGAGCTGA	TAAACATCATG	TTCTCCAATG	TGTTCAGTCGCGCA	8735
Db	8521	CTGGAGCTGA	TAAACATCATG	TTCTCCAATG	TGTTCAGTCGCGCA	8580
QY	8736	AGGATATAC	TACCTACCCGT	GACCCCAAC	CCCTTGGACGCGGCTG	8795
Db	8581	AGGATATAC	TACCTACCCGT	GACCCCAAC	CCCTTGGACGCGGCTG	8640
QY	8796	GCTAGACAC	ACTCCAATCAACT	CTTTCGCTAGGCA	ATATCATGATG	8855
Db	8641	GCTAGACAC	ACTCCAATCAACT	CTTTCGCTAGGCA	ATATCATGATG	8700
QY	8856	TGGCAAGG	ATGATTTCTGAT	GACTCACTTTTCT	TCCATCTTCTAGCT	8915
Db	8701	TGGCAAGG	ATGATTTCTGAT	GACTCACTTTTCT	TCCATCTTCTAGCT	8760
QY	8916	GAAAAAG	CCCTGGATTGTCA			8935
Db	8761	GAAAAAG	CCCTGGATTGTCA			8780

AF054251	8779 bp	RNA	linear	VRL 02-MAY-1998
LOCUS	Hepatitis C virus HC-J4, pCV-J4L1, partial sequence.			
DEFINITION	AF054251			
ACCESSION	AF054251.1			
VERSION	GI:3098640			
KEYWORDS	Hepatitis C virus			
SOURCE	Hepatitis C virus			
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.			
REFERENCE	1 (bases 1 to 8779)			
AUTHORS	Yanagi,M., St Claire,M., Shapiro,M., Emerson,S.U., Purcell,R.H. and Bukh,J.			
TITLE	Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo			
JOURNAL	Virology 244 (1), 161-172 (1998)			
PUBMED	9581788			
REFERENCE	2 (bases 1 to 8779)			
AUTHORS	Yanagi,M. and Bukh,J.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-MAR-1998) Hepatitis Viruses Section, LID, NIAID, NIH, 7 Center Dr., Building 7, Room 201, Bethesda, MD 20892, USA			
FEATURES	Location/Qualifiers			
source	1..8779			
	/organism="Hepatitis C virus"			
	/mol_type="genomic RNA"			
	/strain="HC-J4"			
	/isolate="pCV-J4L1"			
	/db_xref="taxon:11103"			
misc_feature	/note="defective virus"			
	187..>8779			
	/note="region corresponding to polyprotein gene"			
ORIGIN				
Query Match	90.5%;	Score 8684.8;	DB 13;	Length 8779;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 8727;	Conservative	0;	Mismatches 52;	Indels 1; Gaps 1;
QY	156	CGGGTGGAGTACACCGGAATTCGACGACGACCGGGTCTTTCTGGATCAACCCGCTCAA	215	
DB	1	CGGGTGGAGTACACCGGAATTCGACGACGACCGGGTCTTTCTGGATCAACCCGCTCAA	60	
QY	216	TGCCTGGAGATTGGCGGTGCCCGCGAGACTGCTAGCCGAGTAGTGTGGGTGCGCAA	275	
DB	61	TGCCTGGAGATTGGCGGTGCCCGCGAGACTGCTAGCCGAGTAGTGTGGGTGCGCAA	120	
QY	276	AGGCTTGTGGTACTGCTGATAGGGTGTTCGGAGTGCCTGGGAGGTCTCGTAGACCG	335	
DB	121	AGGCTTGTGGTACTGCTGATAGGGTGTTCGGAGTGCCTGGGAGGTCTCGTAGACCG	180	
QY	336	TGCACATGAGCAGCACTAACTCAAGAAACCAACGTAACACCAACCGCGC	395	
DB	181	TGCACATGAGCAGCACTAACTCAAGAAACCAACGTAACACCAACCGCGC	240	
QY	396	CCACAGGAGCTCAAGTTCCCGGCGGTGTGATGCTGTGGTGGAGTTTACCTGTTCGG	455	
DB	241	CCACAGGAGCTCAAGTTCCCGGCGGTGTGATGCTGTGGTGGAGTTTACCTGTTCGG	300	
QY	456	CGCAGGGGCCAGGTTGGGTGTGCGCGACTAGGAAGGCTTCCGAGCGTCCGAACCT	515	
DB	301	CGCAGGGGCCAGGTTGGGTGTGCGCGACTAGGAAGGCTTCCGAGCGTCCGAACCT	360	
QY	516	CGTGAAGCGCACACCTATCCAAAGGCTCCCGACCGGAGGCGGCTGGGCTCAG	575	
DB	361	CGTGAAGCGCACACCTATCCAAAGGCTCCCGACCGGAGGCGGCTGGGCTCAG	420	
QY	576	CCCGGTACCTTGGCCCTCTATGCAATGAGGCGCTGGGTGGGCGAGGATGGCTCCTG	635	
DB	421	CCCGGTACCTTGGCCCTCTATGCAATGAGGCGCTGGGTGGGCGAGGATGGCTCCTG	480	
QY	636	TCACCCCGGGTCCCGGCTAGTTGGGGCCCAACGGAACCCCGGCGTAGGTTCGGTAAAC	695	
DB	481	TCACCCCGGGTCCCGGCTAGTTGGGGCCCAACGGAACCCCGGCGTAGGTTCGGTAAAC	540	
QY	696	TTGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCCGATCTCATGGGGTACATTCG	755	
DB	541	TTGGTAAAGTCAATCGATACCTTACATGCGGCTTCGCCGATCTCATGGGGTACATTCG	600	
QY	756	CTCGTCGGCGCCCTAGGGGGCGCTGCACGGGCTTGGACACGGTGTCCGGGTTCTG	815	
DB	601	CTCGTCGGCGCCCTAGGGGGCGCTGCACGGGCTTGGACACGGTGTCCGGGTTCTG	660	
QY	816	GAGACGGCGTGAATATGCAACAGGAACTTGGCCGTTGCTCTTCTCTATCTTCCTC	875	
DB	661	GAGACGGCGTGAATATGCAACAGGAACTTGGCCGTTGCTCTTCTCTATCTTCCTC	720	
QY	876	TTGGCTCTGCTGCTCTGTTTGACCATCCAGCTTCCGCTTATGAAGTGCACAGCTGCC	935	
DB	721	TTGGCTCTGCTGCTCTGTTTGACCATCCAGCTTCCGCTTATGAAGTGCACAGCTGCC	780	
QY	936	GGGATATACATGTACAGACGACTGCTCAACTCAAGCATGTGTATGAGGACGCGAC	995	
DB	781	GGGATATACATGTACAGACGACTGCTCAACTCAAGCATGTGTATGAGGACGCGAC	840	
QY	996	GTGATCATGATCTCCCGGGTGCCTGCTGTTCAGGAGGGTAAACAGCTCCCGTTGC	1055	
DB	841	GTGATCATGATCTCCCGGGTGCCTGCTGTTCAGGAGGGTAAACAGCTCCCGTTGC	900	
QY	1056	TGGGTAGCGCTCACTCCACCGCTCCCGGCCAGGAATGCCAGCGTCCCGCTACGACAATA	1115	
DB	901	TGGGTAGCGCTCACTCCACCGCTCCCGGCCAGGAATGCCAGCGTCCCGCTACGACAATA	960	
QY	1116	CGACGCACTGCTGCTGCTGTTGGGACGCGCTGTTCTGCTGCTATGATGAGTGGG	1175	
DB	961	CGACGCACTGCTGCTGCTGTTGGGACGCGCTGTTCTGCTGCTATGATGAGTGGG	1020	
QY	1176	GATCTGCGGATCTATTTTCTGCTCCAGCTGTTCACTTCTGCTCCGCGGAT	1235	
DB	1021	GATCTGCGGATCTATTTTCTGCTCCAGCTGTTCACTTCTGCTCCGCGGAT	1080	
QY	1236	GAGACAGTGCAGGACTGCAACTGCTATATCCCGGCCATGTATCAGGTCAACGCGATG	1295	
DB	1081	GAGACAGTGCAGGACTGCAACTGCTATATCCCGGCCATGTATCAGGTCAACGCGATG	1140	
QY	1296	GCTTGGGATATGATGAACTGTGCTACCAAGCCCTAGTGTGTGCGAGTTGCTC	1355	
DB	1141	GCTTGGGATATGATGAACTGTGCTACCAAGCCCTAGTGTGTGCGAGTTGCTC	1200	
QY	1356	CGGATCCCAAGCTGTGCTGCACTGTGGGGGGGCGGCTGCGGAGTCTGCGCGGC	1415	
DB	1201	CGGATCCCAAGCTGTGCTGCACTGTGGGGGGGCGGCTGCGGAGTCTGCGCGGC	1260	
QY	1416	CTTGCTACTATTCCATGTTAGGGAATGCGGCTAAGGTTCTGATTGTGGCGCTACTCTTT	1475	
DB	1261	CTTGCTACTATTCCATGTTAGGGAATGCGGCTAAGGTTCTGATTGTGGCGCTACTCTTT	1320	
QY	1476	GCCGCGTTGACGGGGAGACCAACAGACGGGAGGTGGCGCGGCAACACACCTCCGG	1535	
DB	1321	GCCGCGTTGACGGGGAGACCAACAGACGGGAGGTGGTGGCGGCAACACACCTCCGG	1380	
QY	1536	TTACGTCCTCTTCTCATCTGGGGCGTCTCAGAAATCCAGCTTGTGAATACCAACGCG	1595	
DB	1381	TTACGTCCTCTTCTCATCTGGGGCGTCTCAGAAATCCAGCTTGTGAATACCAACGCG	1440	
QY	1596	AGCTGGCACATCAACAGGACTGCCCTAAATTTGCAATGACTCCCTCCAACTGGGTCTCTTT	1655	
DB	1441	AGCTGGCACATCAACAGGACTGCCCTAAATTTGCAATGACTCCCTCCAACTGGGTCTCTTT	1500	
QY	1656	GCCGCGTTGTTTACGCAACAAGTTCAACTCGTCCGGGTGCCGAGCGCATGCGCAGC	1715	
DB	1501	GCCGCGTTGTTTACGCAACAAGTTCAACTCGTCCGGGTGCCGAGCGCATGCGCAGC	1560	
QY	1716	TGCGCGCCCATGACTGTTGCGCCAGGGGTGGGGCCCATCATCTATCTAAGCTTAAC	1775	
DB	1561	TGCGCGCCCATGACTGTTGCGCCAGGGGTGGGGCCCATCATCTATCTAAGCTTAAC	1620	

QY 1776 AGCTCGGATCAGAGCGCTTATTGCTGGCAATTACGCGCTCGACCGGTGGTGTGCTACCC 1835
DB 1621 AGCTCGGATCAGAGCGCTTATTGCTGGCAATTACGCGCTCGACCGGTGGTGTGCTACCC 1680
QY 1836 GCGTCCAGGTGTGCTCCAGTGATTGTTTCAACCCCAAGCCCTGTGTGGTGGGACC 1895
DB 1681 GCGTCCAGGTGTGCTCCAGTGATTGTTTCAACCCCAAGCCCTGTGTGGTGGGACC 1740
QY 1896 ACCGATCGTTCCGGTGTCCCTACGTATAGCTGGGGGAGAAATGAGACAGACGTGATGCTC 1955
DB 1741 ACCGATCGTTCCGGTGTCCCTACGTATAGCTGGGGGAGAAATGAGACAGACGTGATGCTC 1800
QY 1956 CTCAACAAACAGCGCTCGGCCACAAAGGCAACTGTTTGGCTGTACATGGATGAATAGTACT 2015
DB 1801 CTCAACAAACAGCGCTCGGCCACAAAGGCAACTGTTTGGCTGTACATGGATGAATAGTACT 1860
QY 2016 GGGTTCACTAAGACGTGGGAGGTCCCGGTGTAACATCGGGGGGTGCGTAAACCGACC 2075
DB 1861 GGGTTCACTAAGACGTGGGAGGCCCGCGTGAACATCGGGGGGTGCGTAAACCGACC 1920
QY 2076 TTGATCTGCCCCACGGACTGCTTCCGGAAGCACCCGAGGCTACTTACACAAAATGTGGC 2135
DB 1921 TTGATCTGCCCCACGGACTGCTTCCGGAAGCACCCGAGGCTACTTACACAAAATGTGGC 1980
QY 2136 TCGGGGCCCTGGTTGACACCTAGTGCTTAGTAGACTACCCATACAGGCTTTGGCACTAC 2195
DB 1981 TCGGGGCCCTGGTTGACACCTAGTGCTTAGTAGACTACCCATACAGGCTCTGGCACTAC 2040
QY 2196 CCTGGCACTCTCAATTTTCCATCTTTAAGTTAGGATGATGTGGGGCGTGGAGCAC 2255
DB 2041 CCTGGCACTCTCAATTTTCCATCTTTAAGTTAGGATGATGTGGGGCGTGGAGCAC 2100
QY 2256 AGGCTCAATCCGCAATGGACTCGAGGAGCGCTGTAACTTTGGAGGACAGGGAT 2315
DB 2101 AGGCTCAATCCGCAATGGACTCGAGGAGCGCTGTAACTTTGGAGGACAGGGAT 2160
QY 2316 AGGTCAGAACTACGCCGTGCTGTCTCTAACAGAGTGGCAGATACCTGCCCTGTGCT 2375
DB 2161 AGGTCAGAACTACGCCGTGCTGTCTCTAACAGAGTGGCAGATACCTGCCCTGTGCT 2220
QY 2376 TTCAACACCTACCGGCTTTATCCACTGGTTGATCCATCTCCATCAGAACATCGTGAC 2435
DB 2221 TTCAACACCTACCGGCTTTATCCACTGGTTGATCCATCTCCATCAGAACATCGTGAC 2280
QY 2436 GTGCAATACCTGTACGSGTAGGGTCAGCGTTTCTCTCTTGGCAATCAAATGGAGTAC 2495
DB 2281 GTGCAATACCTGTACGSGTAGGGTCAGCGTTTCTCTCTTGGCAATCAAATGGAGTAC 2340
QY 2496 ATCTGTGTGCTTTTCTCTCTGGCAGACGCGCGGTGTGTGCTCTGTGGATGATG 2555
DB 2341 ATCTGTGTGCTTTTCTCTGGCAGACGCGCGGTGTGTGCTCTGTGGATGATG 2400
QY 2556 CTGCTGATAGCCCAAGGCTAGGCGCGCTTAGAGAACTTGGTGGTCCCTCAATGGCGGCTCC 2615
DB 2401 CTGCTGATAGCCCAAGGCTAGGCGCGCTTAGAGAACTTGGTGGTCCCTCAATGGCGGCTCC 2460
QY 2616 GTGGCCGAGCGCATGTTCTCTCTCTTCTGTGTTCTTCTCGCGCGCTGTGACATT 2675
DB 2461 GTGGCCGAGCGCATGTTCTCTCTCTTCTGTGTTCTTCTCGCGCGCTGTGACATT 2520
QY 2676 AAGGCGAGGCTGGCTCTCTGGGCGCGGCTPATGCTTTTATGGCGTATGGCCGCTCCTG 2735
DB 2521 AAGGCGAGGCTGGCTCTCTGGGCGCGGCTPATGCTTTTATGGCGTATGGCCGCTCCTG 2580
QY 2736 CTCTACTGGCGTTACCAACAGAGCTTACGCTTTGGACCGGGAGATGGCTGCACTGTCG 2795
DB 2581 CTCTACTGGCGTTACCAACAGAGCTTACGCTTTGGACCGGGAGATGGCTGCACTGTCG 2640
QY 2796 GGGGGTGGCTCTGTGAGTCTGGTATTCTTGACCTTGTACCATACTCAAAAGTGTTT 2855
DB 2641 GGGGGTGGCTCTGTGAGTCTGGTATTCTTGACCTTGTACCATACTCAAAAGTGTTT 2700
QY 2856 CTCACTAGGCTCATATGGTGGTTACAAATACATTTTATCACAGAGCGGCGCACATGCAA 2915

DB 2701 CTCACTAGGCTCATATGGTGGTTACATACTTTATCACAGAGCCGAGGGGCACATGCAA 2760
QY 2916 GTGTGGTCCCCCCTCAACGTTTGGGAGGCGCGATGCCCATCATCTCTCTCAGTGT 2975
DB 2761 GTGTGGTCCCCCCTCAACGTTTGGGAGGCGCGATGCCCATCATCTCTCTCAGTGT 2820
QY 2976 GGGTTTCATCCAGAGTTAATTTTTCACATCACCAAACTCCTGCTGCCATACCTCGGCCG 3035
DB 2821 GGGTTTCATCCAGAGTTAATTTTTCACATCACCAAACTCCTGCTGCCATACCTCGGCCG 2880
QY 3036 CTCATGGTCTCCAGGCTGGCATAACGAGGTGCCGTACTTTCGTGGCGCTCAAGGGCTC 3095
DB 2881 CTCATGGTCTCCAGGCTGGCATAACGAGGTGCCGTACTTTCGTGGCGCTCAAGGGCTC 2940
QY 3096 ATTCTGTGATGATGTTAGTGCAGAAAGTCGCGGGGTCAATTATGTCAAAATGTCCTTC 3155
DB 2941 ATTCTGTGATGATGTTAGTGCAGAAAGTCGCGGGGTCAATTATATCAAAATGGGCTTC 3000
QY 3156 ATGAAGCTGGCGCGCTGACAGGTAGTACGTTTATAACCACTTTACCCACTGCGGGAC 3215
DB 3001 ATGAAGCTGGCGCGCTGACAGGTAGTACGTTTATAACCACTTTACCCACTGCGGGAC 3060
QY 3216 TGGGCCACGCGCGCTTACGAGACCTTTCGGTGGCGTAGAGCCGTGCTCTTCGCGC 3275
DB 3061 TGGGCCACGCGCGCTTACGAGACCTTTCGGTGGCGTAGAGCCGTGCTCTTCGCGC 3120
QY 3276 ATGGAGACCAAGGTATCATACCTGGGAGCAGACACCGCTGCTGTGGGACATCATCTTG 3335
DB 3121 ATGGAGACCAAGGTATCATACCTGGGAGCAGACACCGCTGCTGTGGGACATCATCTTG 3180
QY 3336 GGTCTACCGTCTCCGCCGAGGGGAGGAGATATTTTGGGACCGGCTGATGCTC 3395
DB 3181 GGTCTACCGTCTCCGCCGAGGGGAGGAGATATTTTGGGACCGGCTGATGCTC 3240
QY 3396 GAAGGCAAGGGTGGGACCTCTTCGCGCCATCACGCGCTACTCCCAAAACGCGGGC 3455
DB 3241 GAAGGCAAGGGTGGGACCTCTTCGCGCCATCACGCGCTACTCCCAAAACGCGGGC 3300
QY 3456 GTACTTGGTTGATCATACCTAGCCTCACAGCGGGACAAAGAACAGGTGCAAGGGAG 3515
DB 3301 GTACTTGGTTGCTCATACCTAGCCTCACAGCGCGGACAAAGAACAGGTGCAAGGGAG 3360
QY 3516 GTTCAAGTGGTTTCTACCGCAACACAACTTTCCTGGCGACCTGCATCAACGGGCTGTC 3575
DB 3361 GTTCAAGTGGTTTCTACCGCAACACAACTTTCCTGGCGACCTGCATCAACGGGCTGTC 3420
QY 3576 TGGACTCTCTACCATGCGCTGCTCGAAGACCTAGCCGCTCCAAAGGTCCAATCACC 3635
DB 3421 TGGACTCTCTACCATGCGCTGCTCGAAGACCTAGCCGCTCCAAAGGTGCC-ATCACC 3479
QY 3636 CAAATGTACACCAATGTAGACCTGGACCTCGTGGCTGGCAGGCGCCCCCGGGGCGCGC 3695
DB 3480 CAAATGTACACCAATGTAGACCTGGACCTCGTGGCTGGCAGGCGCCCCCGGGGCGCGC 3539
QY 3696 TCATGACACCATGCACTGTGGAGCTCGGACCTTTATCTTGTGTACAGACATGCTGAT 3755
DB 3540 TCATGACACCATGCACTGTGGAGCTCGGACCTTTTACTTTGGTTCACGAGACATGCTGAT 3599
QY 3756 GTCAATTCCGTGGCGCGGAGGCGACAGAGGGGAGTCTACTCTCCCGAGGCCGCTC 3815
DB 3600 GTCAATCCGTGGCGCGGAGGCGACAGAGGGGAGTCTACTCTCTCCCGAGGCCGCTC 3659
QY 3816 TCTTACTGAAAGGCTCCTCGGGTGGTCCATTGTTTGGCCCTTCGGGGCAGCTCGTGGGC 3875
DB 3660 TCTTACTGAAAGGCTCCTCGGGTGGTCCATTGTTTGGCCCTTCGGGGCAGCTCGTGGGC 3719
QY 3876 GTCTTCCGGGCTGTGTGACACCGGGGGGTGCGCAAGCGGTGGACTTCATACCGGTT 3935
DB 3720 GTCTTCCGGGCTGTGTGACACCGGGGGGTGCGCAAGCGGTGGACTTCATACCGGTT 3779
QY 3936 GAGTCTATGMAACTACCATGCGGTCTCCGCTCTTTCACAGACAACTCAACCCCGCGCT 3995

Db 3780 GAGTCTATGGAAATPACCATGCGGTCCCGGTCTTTCACAGACAACTCATCCCCCGCGCT 3839
QY 3996 GTACCGCAGACATTCCAAGTGGACATCTGCAAGCTCTCTACTGGCAGCGGCAAGAGCAC 4055
Db 3840 GTACCGCAGACATTCCAAGTGGACATCTGCAAGCTCTCTACTGGCAGCGGCAAGAGCACT 3899
QY 4056 AAAAGTSCGGCTGCGTATGCGACCCAAAGGGTACAAAGTGTCTGCTCTCTGAAACCGGTCCGTT 4115
Db 3900 AAAAGTSCGGCTGCGTATGCGACCCAAAGGGTACAAAGTGTCTGCTCTCTGAAACCGGTCCGTT 3959
QY 4116 GCCGCCACTTATAGGTTTGGGGCGTATATGTCCAAAGGCACAGGTATCGACCTTAACATC 4175
Db 3960 GCCGCCACTTATAGGTTTGGGGCGTATATGTCCAAAGGCACAGGTATCGACCTTAACATC 4019
QY 4176 AGAACTGGGGTAAGACCAATTACCAACGGCGGCTCCATTAGTACTCCACCTATGGCAAG 4235
Db 4020 AGAACTGGGGTAAGACCAATTACCAACGGCGGCTCCATTAGTACTCCACCTATGGCAAG 4079
QY 4236 TTCTCTTGGCGACGGTGGCTGTTCTGGGGCGGCTATGACATCATATAATATGTGATGAGTGC 4295
Db 4080 TTCTCTTGGCGACGGTGGCTGTTCTGGGGCGGCTATGACATCATATAATATGTGATGAGTGC 4139
QY 4296 CACTCAACTGACTCGACTACCATCTTTGGGCATTCGGCAGAGTCTCTGGACAAAGCGGAGACG 4355
Db 4140 CATTCAACTGACTCGACTACCATCTTTGGGCATTCGGCAGAGTCTCTGGACAAAGCGGAGACG 4199
QY 4356 GCTGAGCGCGCTCGTGTCTCGCCACCGCTACACCTCCGGGATCGGTTACCGTGCCA 4415
Db 4200 GTTGGAGCGCGCTGTGTGTCTCGCCACCGCTACACCTCCGGGATCGGTTACCGTGCCA 4259
QY 4416 CACCCCAATATCGAGGAATAGCGCTGTCCAAATGAGAGATCCCTTCTATGGCAAA 4475
Db 4260 CACCCCAATATCGAGGAATAGCGCTGTCCAAATGAGAGATCCCTTCTATGGCAAA 4319
QY 4476 GCATCCCCCATTTAGGCCATCAAGGGGGAGGCACTCTCATTTTCTGCCATTTCCAAAGAG 4535
Db 4320 GCCATCCCCCATTTAGGCCATCAAGGGGGAGGCACTCTCATTTTCTGCCATTTCCAAAGAG 4379
QY 4536 AAATGTAGAGCTCGCGCAAGCTGACAGGCTCGGACTGAAACGCTGTAGCATATTAC 4595
Db 4380 AACTGCGAGAGCTCGCGCAAGCTGACAGGCTCGGACTGAAACGCTGTAGCATATTAC 4439
QY 4596 CCGGGCTTGATGTGTCCGTCTATACCGCTATCGGAGAGCTGTTGTCTGGGCAACAGAC 4655
Db 4440 CCGGGCTTGATGTGTCCGTCTATACCGCTATCGGAGAGCTGTTGTCTGGGCAACAGAC 4499
QY 4656 GCTCTAATGACGGGTTTACCGGCGATTTTGAATCAGTGATCGACTGCAATACATGTGTC 4715
Db 4500 GCTCTAATGACGGGTTTACCGGCGATTTTGAATCAGTGATCGACTGCAATACATGTGTC 4559
QY 4716 ACCCAGACAGTCTGACTTCTGAGCTTGGATCCACCTTCCACCATTTGAGACGACCGGTGCC 4775
Db 4560 ACCCAGACAGTCTGACTTCTGAGCTTGGATCCACCTTCCACCATTTGAGACGACCGGTGCC 4619
QY 4776 CAAAGCGCGGTGTCCGCTCGCAACCGCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATC 4835
Db 4620 CAGGACCGCGGTGTCCGCTCGCAACCGCGAGGTAGAACTGGCAGGGGTAGGAGTGGCATC 4679
QY 4836 TACAGGTTTGTGATCTCAGGAGAACCGCCCTCGGGCATGTTTCAATTTCTCGGTCTGTGT 4895
Db 4680 TACAGGTTTGTGATCTCAGGAGAACCGCCCTCGGGCATGTTTCAATTTCTCGGTCTGTGT 4739
QY 4896 GAGTGTATGACCGGGCTGTGTTGGTATGAGCTCACCGCGCTGAGACCTCGGTTAGG 4955
Db 4740 GAGTGTATGACCGGGCTGTGTTGGTATGAGCTCACCGCGCTGAGACCTCGGTTAGG 4799
QY 4956 TTGCGGGCTTACTTAATACACAGGGTTGCCGCTCTGCGAGACCATCTCGAGTTCTGG 5015
Db 4800 TTGCGGGCTTACTTAATACACAGGGTTGCCGCTCTGCGAGACCATCTCGAGTTCTGG 4859
QY 5016 GAGAGCGTCTTACAGGCTTCAACCAATAGATGCCACTTCTGTGCCAGACTAAACAG 5075
Db 4860 GAGAGCGTCTTACAGGCTTCAACCAATAGATGCCACTTCTGTGCCAGACTAAACAG 4919

QY 5076 GCAGGAGACAACTTTCTTACCTGGTGGCATATCAAGCTACAGTGTGCGCCAGGCTCAA 5135
Db 4920 GCAGGAGACAACTTTCTTACCTGGTGGCATATCAAGCTACAGTGTGCGCCAGGCTCAA 4979
QY 5136 GCTCCACTCATCGTGGGACCAAAATGTGGAAGTGTCTCATACGGCTGAAACCTACACTG 5195
Db 4980 GCTCCACTCATCGTGGGACCAAAATGTGGAAGTGTCTCATACGGCTGAAACCTACACTG 5039
QY 5196 CACGGGCCAACACCCCTGCTGTAGCTAGGAGCGCTCCAAAATAGGTCATCTCACA 5255
Db 5040 CACGGGCCAACACCCCTGCTGTAGCTAGGAGCGCTCCAAAATAGGTCATCTCACA 5099
QY 5256 CACCCCAATAACTAAATACATCATGGCATGTCTGCGTACCTCGAGGTGCTCACTAGC 5315
Db 5100 CACCCCAATAACTAAATACATCATGGCATGTCTGCGTACCTCGAGGTGCTCACTAGC 5159
QY 5316 ACTTGGGTGTGTGAGCGGAGTCTTTGACAGCTTTTGGCCGCATATCTGCTGACGACAGC 5375
Db 5160 ACTTGGGTGTGTGAGCGGAGTCTTTGACAGCTTTTGGCCGCATATCTGCTGACGACAGC 5219
QY 5376 AGTGTGCTCATTTGTGGCAGGATCATCTTGTCCGGGAAGCCAGCTGTCTGTTCCGACAGG 5435
Db 5220 AGTGTGCTCATTTGTGGCAGGATCATCTTGTCCGGGAAGCCAGCTGTCTGTTCCGACAGG 5279
QY 5436 GAAGTCTCTTACACAGGAGTTCGATGAGATGGAAGTGTGCTCACAACCTTCTTACATC 5495
Db 5280 GAAGTCTCTTACACAGGAGTTCGATGAGATGGAAGTGTGCTCACAACCTTCTTACATC 5339
QY 5496 GAGCAGGGAATGCAAGCTCGCGAGCAATTTAAGCAAAAGGCGCTCGGGTTGTGCAACG 5555
Db 5340 GAGCAGGGAATGCAAGCTCGCGAGCAATTTAAGCAAAAGGCGCTCGGGTTGTGCAACG 5399
QY 5556 GCCACCAAGCAAGCGGAGGCTGCTCCCGTGTGGAGTCCAAAGTGGCGAGCCCTTCAG 5615
Db 5400 GCCACCAAGCAAGCGGAGGCTGCTCCCGTGTGGAGTCCAAAGTGGCGAGCCCTTCAG 5459
QY 5616 ACCTTCTGGGGCAAGCAGCATGTGGAATTTTCATCAGCGGAATACAGTACTTAGCAGGCTTA 5675
Db 5460 ACCTTCTGGGGCAAGCAGCATGTGGAATTTTCATCAGCGGAATACAGTACTTAGCAGGCTTA 5519
QY 5676 TCCACTCTGCTGGAACCCCGCGATAGCATCATTTGATGGCATTTACAGTCTTATCACT 5735
Db 5520 TCCACTCTGCTGGAACCCCGCGATAGCATCATTTGATGGCTTTTACAGTCTTATCACT 5579
QY 5736 AGCCGCTCACACCCCAAAACACCTCTCTGTTTAAACATCTTGGGGGATGGGTGCTGCC 5795
Db 5580 AGCCGCTCACACCCCAAAACACACCTCTCTGTTTAAACATCTTGGGGGATGGGTGCTGCC 5639
QY 5796 CAACTCGCTCTCCAGCGCTGCGTACGCTTTTCTGTCGCGCGCATCGCCGAGCGGCT 5855
Db 5640 CAACTCGCTCTCCAGCGCTGCGTACGCTTTTCTGTCGCGCGCATCGCCGAGCGGCT 5699
QY 5856 GTTGGCAGCATAGGCTTTGGGAAAGGTGCTCTGTGACATCTTTGGGGGCTATGGGCGAGG 5915
Db 5700 GTTGGCAGCATAGGCTTTGGGAAAGGTGCTCTGTGACATCTTTGGGGGCTATGGGCGAGG 5759
QY 5916 GTAGCCGCGCATCTGCTGGCTTTAAGGTCTAGCGCGGAGGTGCCCTCCACCGAGGAC 5975
Db 5760 GTAGCCGCGCATCTGCTGGCTTTAAGGTCTAGCGCGGAGGTGCCCTCCACCGAGGAC 5819
QY 5976 CTGGTCAACTTACTCCCTGCCATCTCTCTCTGTCGCTGCTGCTGCGGCTCGGTGTC 6035
Db 5820 CTGGTCAACTTACTCCCTGCCATCTCTCTCTCTGTCGCTGCTGCTGCGGCTCGGTGTC 5879
QY 6036 GCAGCAATACTCGGTGCGCACGTGGGCCCGGAGAGGGGCTGTGACGTGAGTAAACCGG 6095
Db 5880 GCAGCAATACTCGGTGCGCACGTGGGCCCGGAGAGGGGCTGTGACGTGAGTAAACCGG 5939
QY 6096 CTGATAGGCTTCTGCTTCCGGGGTAAACACGCTCTCCCTACGCACTATGTGCTGAGC 6155
Db 5940 CTGATAGGCTTCTGCTTCCGGGGTAAACACGCTCTCCCTACGCACTATGTGCTGAGAC 5999

QY 6156 GACGCTGCAGCAGCTGTCTCACTCAGATCCCTCTCTAGCCTTACATCACTCAACTGCTGAAG 6215
Db 6000 GACGCTGCAGCAGCTGTCTCACTCAGATCCCTCTCTAGCCTTACATCACTCAACTGCTGAAG 6059
QY 6216 CGGCTCCACAGTGAATTAATGAGGACTGCTCTACGCCATGCTCCGGCTCGTGGCTAAGS 6275
Db 6060 CGGCTTCACAGTGGATTAATGAGGACTGCTCTACGCCATGCTCCGGCTCGTGGCTAAGS 6119
QY 6276 GATGTTTGGGATTTGGATATGACAGGTTGACTGACTTCAAGACCTGGCTCCAGTCCAAA 6335
Db 6120 GATGTTTGGGATTTGGATATGACAGGTTGACTGACTTCAAGACCTGGCTCCAGTCCAA 6179
QY 6336 CTCTGCGCGGTATACGGGAGTCCCTTCTCTGCTCATGCAACCGCGGTACAGGAGTC 6395
Db 6180 CTCTGCGCGGTATACGGGAGTCCCTTCTCTGCTCATGCAACCGCGGTACAGGAGTC 6239
QY 6396 TGGCGGGGGACGGCATCATGCAAAACCACTGCCATGCGGAGCAGAGATCGCGGACAT 6455
Db 6240 TGGCGGGGGACGGCATCATGCAAAACCACTGCCATGCGGAGCAGAGATCGCGGACAT 6299
QY 6456 GTCAAAAACGGTTCCATGAGGATCGTAGGCTTAGAACCTGCAGCAACACGTTGGCACGGA 6515
Db 6300 GTCAAAAACGGTTCCATGAGGATCGTAGGCTTAGAACCTGCAGCAACACGTTGGCACGGA 6359
QY 6516 ACCTTCCCATCAACGATACACACGGGACCTTTGCACACCTCCCGGCGCCCAACTAT 6575
Db 6360 ACCTTCCCATCAACGATACACACGGGACCTTTGCACACCTCCCGGCGCCCAACTAT 6419
QY 6576 TCAGGGCGCTATGGCGGTGGCTGTGAGGAGTACGTTGAGGTTACGCTGTGGGGAT 6635
Db 6420 TCAGGGCGCTATGGCGGTGGCTGTGAGGAGTACGTTGAGGTTACGCTGTGGGGAT 6479
QY 6636 TTCCACTACGTGACGGGATGACCACTGACAAAGTGCACAGTCCAGAGTTCCGGCC 6695
Db 6480 TTCCACTACGTGACGGGATGACCACTGACAAAGTGCACAGTCCAGAGTTCCGGCC 6539
QY 6696 CCCGAATTTCTTCAGGAGTGGATGAGTGGGTTGCACAGGTACGCTCCGGGTCGAAA 6755
Db 6540 CCCGAATTTCTTCAGGAGTGGATGAGTGGGTTGCACAGGTACGCTCCGGGTCGAAA 6599
QY 6756 CTTCTTCTACGGGAGGACGTCACTTCAGGTCCGGTCCAAACCAATCTTGGTGGGTGCG 6815
Db 6600 CTTCTTCTACGGGAGGACGTCACTTCAGGTCCGGTCCAAACCAATCTTGGTGGGTGCG 6659
QY 6816 CAGTCCCATCGAGCCGACCGGACGTACAGTGTCTTCTTCCATGCTCACCGATCCC 6875
Db 6660 CAGTCCCATCGAGCCGACCGGACGTAAACAGTGTCTTCTTCCATGCTCACCGATCCC 6719
QY 6876 TCCACATTAACAGCAGAGACGGCTAAGCTAGGCTGCTAGAGGTCTCCCGCTCTTTA 6935
Db 6720 TCCACATTAACAGCAGAGACGGCTAAGCTAGGCTGCTAGAGGTCTCCCGCTCTTTA 6779
QY 6936 GCCAGCTCATCAGCTAGCAGTGTCTGCGCTTTCTTTGAGGCGACATGACATCCAC 6995
Db 6780 GCCAGCTCATCAGCTAGCAGTGTCTGCGCTTTCTTTGAGGCGACATGACATCCAC 6839
QY 6996 CATGACTCCCGGACGCTGACCTCATGAGGCGCAACCTCTTGTGGCGGAGAGATGGC 7055
Db 6840 CATGACTCCCGGACGCTGACCTCATGAGGCGCAACCTCTTGTGGCGGAGAGATGGGT 6899
QY 7056 GGAACATCATCTCGGTGGAGTCAGAGNATAGGTAGTAAATCTTGGACTCTTTTGAACCG 7115
Db 6900 GGAACATCATCTCGGTGGAGTCAGAGNATAGGTAGTAAATCTTGGACTCTTTTGAACCG 6959
QY 7116 CTTCCAGCGGAGGGGATGAGAGGAGATATCCGTCGCGGCGGAGATCCTCGCAAAATCC 7175
Db 6960 CTTCCAGCGGAGGGGATGAGAGGAGATATCCGTCGCGGCGGAGATCCTCGCAAAATCC 7019
QY 7176 AGGAAGTTCCCGCTCAGCGTTCCCATATGGGACCGCCCGGACTCAATCTCTCACTGCTA 7235
Db 7020 AGGAAGTTCCCGCTCAGCGTTCCCATATGGGACCGCCCGGACTCAATCTCTCACTGCTA 7079
QY 7236 GAGTCTTGGAAGGACCCGGACTACGCTCCCTCGGTGGTACAGGATGCCCATTTGCCACCT 7295

Db 7080 GAGTCTTGGAAGACCCGGACTACGTCCTCCGTTGGTACACGGATGCCATTTGCCACCT 7139
QY 7296 ACCAAGGCTCTCTCCAAATACACCTCCACGGAGAAAAGAGACGGTTGTCTTGACAGAAATCC 7355
Db 7140 ACCAAGGCTCTCTCCAAATACACCTCCACGGAGAAAAGAGACGGTTGTCTTGACAGAAATCC 7199
QY 7356 AATGTGTCTTCTGCTTGGCGGAGCTCCCACTTAGACCTTCGGTAGCTCCGGATCGTCG 7415
Db 7200 AATGTGTCTTCTGCTTGGCGGAGCTCCCACTTAGACCTTCGGTAGCTCCGGATCGTCG 7259
QY 7416 GCGTGTGATAGCGGACCGGCGACCGCCCTCTCTGACCTGGCTCCGACGACGGTGACAAA 7475
Db 7260 GCGTGTGATAGCGGACCGGCGACCGCCCTCTCTGACCTGGCTCCGACGACGGTGACAAA 7319
QY 7476 GGATCCGACGTTGAGTGTGCTCTCTCATGTCGCCCTTTGAAAGGGAGCGCGGGGACCCC 7535
Db 7320 GGGTCCGACGTTGAGTGTGCTCTCTCATGTCGCCCTTTGAAAGGGAGCGCGGGGACCCC 7379
QY 7536 GATCTCAGCGACGGGTCTTGGTCTACCGTGTGAGGAGGCTAGTGAGGATGTCTGCTGC 7595
Db 7380 GATCTCAGCGACGGGTCTTGGTCTACCGTGTGAGGAGGCTAGTGAGGATGTCTGCTGC 7439
QY 7596 TGTCTCAATCTCTATACGTGGACAGCGCCCTGATCAGCCATGCGCTGCGGAGGAAAGT 7655
Db 7440 TGTCTCAATCTCTATACGTGGACAGCGCCCTGATCAGCCATGCGCTGCGGAGGAAAGT 7499
QY 7656 AAGCTCCCATCAACCCGTTGAGCAACTCTTGTCTGCTGTCACCAACAATGCTCTACGCC 7715
Db 7500 AAGCTCCCATCAACCCGTTGAGCAACTCTTGTCTGCTGTCACCAACAATGCTCTACGCC 7559
QY 7716 ACAACATCCCGCAGCGCAAGCCTCCGGCAGAAAGGTACCTTTTGACAGATTGCAAGTC 7775
Db 7560 ACAACATCCCGCAGCGCAAGCCTCCGGCAGAAAGGTACCTTTTGACAGATTGCAAGTC 7619
QY 7776 CTGGATGATCATTTACCGGACGCTACTCAAGGAGATGAGCGGAGGCTCCACAGTTAAG 7835
Db 7620 CTGGATGATCATTTACCGGACGCTACTCAAGGAGATGAGCGGAGGCTCCACAGTTAAG 7679
QY 7836 GCTAAGCTTCTATCTATAGAGGAGGCTGCAAGCTGACGCCCCCAATTTCCGCAAAATCC 7895
Db 7680 GCTAAGCTTCTATCTATAGAGGAGGCTGCAAGCTGACGCCCCCAATTTCCGCAAAATCC 7739
QY 7896 AATTTGGCTATGGGCAAGGACGTCGGAACCTATCCAGAGGGCGCTTTAACCAATC 7955
Db 7740 AATTTGGCTATGGGCAAGGACGTCGGAACCTATCCAGAGGCGCTTTAACCAATC 7799
QY 7956 CGTCTCGTGTGGGAGGACTTGTGGAGACACTGAAACCAATTTGACACCAATCATG 8015
Db 7800 CGTCTCGTGTGGGAGGACTTGTGGAGACACTGAAACCAATTTGACACCAATCATG 7859
QY 8016 GCAAAAGTGAAGTTTCTCGCTCCACAGAGAGGGAGCGCGCAAGCCAGCTCGGCTT 8075
Db 7860 GCAAAAGTGAAGTTTCTCGCTCCACAGAGAGGGAGCGCGCAAGCCAGCTCGGCTT 7919
QY 8076 ATCGTATTTCCAGACCTGGGAGTTCTGTATGCGAGAGATGGCCCTTTACGAGTGGTC 8135
Db 7920 ATCGTATTTCCAGACCTGGGAGTTCTGTATGCGAGAGATGGCCCTTTACGAGTGGTC 7979
QY 8136 TCCACCTCTCTCAGGCGGTGATGGGCTCTCATACGGAATTTCAATATCTCCCCCAAGCAG 8195
Db 7980 TCCACCTCTCTCAGGCGGTGATGGGCTCTCATACGGAATTTCAATATCTCCCCCAAGCAG 8039
QY 8196 CGGTCGAGTCTCTGTGTAATCCTGGAATCAAGAAATGCCCTATGGGCTTCTCATAT 8255
Db 8040 CGGTCGAGTCTCTGTGTAATCCTGGAATCAAGAAATGCCCTATGGGCTTCTCATAT 8099
QY 8256 GACACCGCTGTTTGTGACTCAACGCTCAGAGTGCATCTCGTGTGTGAGGAGTCAAT 8315
Db 8100 GACACCGCTGTTTGTGACTCAACGCTCAGAGTGCATCTCGTGTGTGAGGAGTCAAT 8159
QY 8316 TACCAATGTTGTACTTGGCCCCCGGAGCCAGACAGGCCAATAAGGTGCTCTCAGAGCGG 8375

[illegible]

1201	CGGATCCACAAAGCTGCTCGTGGACATGGTGGCGGGGCCCACTGGGGAGTCTCTGGCGGG	1260
1415	CTTTGGCTACTATTCCATCGGTAGGGAACCTGGGCTAAGGTTTGTGATTTGGCGGCTACTCTTT	1474
1261	CTTTGGCTACTATTCCATCGGTAGGGAACCTGGGCTAAGGTTTGTGATTTGGCGGCTACTCTTT	1320
1475	TGCGGGGTTGACGGGGAGACCCACACAGACGGGAGGGTGGCGGGCCACACACCACTCCGG	1534
1321	TGCGGGGTTGACGGGGAGACCCACACAGACGGGAGGGTGGTGGCGGACCACTCCCGG	1380
1535	GTTCACGTCCTTTTCTCATCTGGGCGTCTCAGAAAAATCCAGCTTGTGAATACCAACGG	1594
1381	GTTTACGTCCTTTTCTCATCTGGGCGTCTCAGAAAAATCCAGCTTGTGAATACCAACGG	1440
1595	CAGCTGGCACAACACAGACTGCCCTTAAATTTGCANTGACTCCCTCCAACTTGGGTTCTT	1654
1441	CAGCTGGCACAACACAGGACTGCTCTTAAATTTGCAATGACTCCCTCCAGACTGGGTTCTT	1500
1655	TGCGGGCTGTTTTTACGCAACACAAGTTTCAACTCGTCCGGGTGCCGAGCGCATGCCACG	1714
1501	TGCGGGCTGTTTTTACACACANAAGTTTCAACTCGCCGGGTGCCGAGGGCATGGCCAG	1560
1715	CTCGCGCCCATTTGACTGGTTCCGCCAGGGTGGGCCCCATCATCCTATACTAAGGCTTAA	1774
1561	CTCGCGCCCATTTGACTGGTTCCGCCAGGGTGGGCCCCATCACCCTACTAGGCTTAA	1620
1775	CAGCTGGATCAGAGGCTTATTGCTGGCATTAACGCGCTCGACCGTGTGGTGTCTCTACC	1834
1621	CAGCTGGATCAGAGGCTTATTGCTGGCATTAACGCGCTCGACCGTGTGGTATCTGTACC	1680
1835	CGGTCGCAAGTGTGGTCCAGTGATTGTTTTCACCCAAAGCCCTGTTGTGGTGGGAC	1894
1681	CGGTCGCAAGTGTGGTCCAGTGATTGTTTTCACCCAAAGCCCTGTTGTGGTGGGAC	1740
1895	CACCGATGTTCCGGTGTCCCTACGTATAGCTGGGGGAGAAATGAGACAGACTGATGCT	1954
1741	CACCGATGTTCCGGTGTCCCTACGTATAGCTGGGGGAGAAATGAGACAGACTGATGCT	1800
1955	CCTCAACAAACAGCGTCCGCCACAAGGCAACTGGTTCCGGCTGTACATGATGAATAGTAC	2014
1801	CCTGAAACAACAGCGTCCGCCACAAGGCAACTGGTTCCGGCTGTACATGATGAATAGTAC	1860
2015	TGGGTTCACTAAGACGTGGGAGGTCCCCGTGTAAATCGGGGGGTCCGTAACCGCAC	2074
1861	TGGGTTCACTAAGACGTGGGAGGCCCCGTGTAAATCGGGGGGTCCGTAACCGCAC	1920
2075	CTTCATCTGCCCCACGACTGCTTCCGGAAGCACCCCGAGGCTACTTACACAAAATGTGG	2134
1921	CTTGATCTGCCCCACGACTGCTTCCGGAAGCACCCCGAGGCTACTTACACAAAATGTGG	1980
2135	CTCGGGCCCTGTTGACACTAGTGGCTTAGTAGACTACCAATACAGGCTTTGGGACTA	2194
1981	CTCGGGCCCTGTTGACACTAGTGGCTTAGTAGACTACCAATACAGGCTTTGGGACTA	2040
2195	CCCTGCACTCTCAATTTTTCATCTTTAGGTTAGGATGTATGTGGGGGGCGTGGAGCA	2254
2041	CCCTGCACTCTCAATTTTTCATCTTTAGGTTAGGATGTATGTGGGGGGCGTGGAGCA	2100
2255	CAGGCTCAATGCCCATGCAATTTGGAATCTCGAGGAGAGCGCTGTAACTTGGAGGACAGGGA	2314
2101	CAGGCTCAATGCCCATGCAATTTGGAATCTCGAGGAGAGCGCTGTAACTTGGAGGACAGGGA	2160
2315	TAGGTCAGAACTCAGCCCGCTGCTGTCTTACAACAGAGTGGGAGATCTGCCCTGTGC	2374
2161	TAGGTCAGAACTCAGCCCGCTGCTGTCTTACAACAGAGTGGGAGATCTGCCCTGTGC	2220
2375	TTTCAACACCTTACGGGCTTTATCCACTGGTTTGATCCATCTCCATCAGAACATCTGGGA	2434
2221	CTTCAACACCTTACGGGCTTTATCCACTGGTTTGATCCATCTCCATCAGAACATCTGGGA	2280
2435	CGTGAATACCTGTACGGGTAGGGTCAGCGTTTGTCTCTTTTGCATTCAAATGGGAGTA	2494
2281	CGTGAATACCTGTACGGGTAGGGTCAGCGTTTGTCTCTTTTGCATTCAAATGGGAGTA	2340

QY 2495 CATCTGTTGCTTTTCCCTTCTCCTGGCAGACGCGCGTGTGTGCTTGTGTGATGAT 2554
DB CATCTGTTGCTTTTCCCTTCTCCTGGCAGACGCGCGAGTGTGTGCTTGTGTGATGAT 2400
QY 2555 GCTGCTGATAGCCCAAGGCTGAGGCGCCTTAGAGAACTTGTGTGTCCTCAATGCGGGGTC 2614
DB GCTGCTGATAGCCCAAGGCTGAGGCGCCTTAGAGAACTTGTGTGTCCTCAATGCGGGGTC 2460
QY 2615 CGTGGCCGAGGGGATGATATCTCTCTCTTCTTGTGTGTCTTCTGCGCGCGCTGGTACAT 2674
DB CGTGGCCGAGGGGATGATATCTCTCTCTTCTTGTGTGTCTTCTGCGCGCGCTGGTACAT 2520
QY 2675 TAAGGGCAGGCTGGCTCCTGGGGCGGCTATGTCTTTTATGTGGCGTATGGCGGCTGTCTCT 2734
DB TAAGGGCAGGCTGGCTCCTGGGGCGGCGTATGTCTTTTATGTGGCGTATGGCGGCTGTCTCT 2580
QY 2735 GCTCTACTGGCGGTTACCACCAAGCTTACGCTTTGACCGGGAGATGGCTGCATCGTG 2794
DB GCTCTACTGGCGGTTACCACCAAGCTTACGCTTTGACCGGGAGATGGCTGCATCGTG 2640
QY 2795 CGGGGGTGGGTTCTTGTAGTCTGGTATTTCTTGACCTTGTCAACATACATAAAAGTGT 2854
DB CGGGGGTGGGTTCTTGTAGTCTGGTATTTCTTGACCTTATCACCATATCTCAAAGTGT 2700
QY 2855 TCTCACTAGGCTCATATGGTGGTTACAATACCTTTATCACAGAGCGGAGCGCATGCA 2914
DB TCTCACTAGGCTCATATGGTGGTTACAATACCTTTATCACAGAGCGGAGCGCATGCA 2760
QY 2915 AGTGTGGGTCCCGCCCTCAAGCTTCGGGGAGCGCGATGCATCATCTCTCTCACGTG 2974
DB AGTGTGGGTCCCGCCCTCAAGCTTCGGGGAGCGCGATGCATCATCTCTCTCACGTG 2820
QY 2975 TCGCGTTCAATCAGAGTTAAATTTTGTGACATACCAAACTCTCTGTGCGGCATCTCGGCC 3034
DB TCGCGTTCAATCAGAGTTAAATTTTGTGACATACCAAACTCTCTGTGCGGCATCTCGGCC 2880
QY 3035 GCTCATGTGTCCAGGCTGGCATACGAGAGTCCGTACTTCTGTCGGCTCAAGGCT 3094
DB GCTCATGTGTCCAGGCTGGCATACGAGAGTCCGTACTTCTGTCGGCTCAAGGCT 2940
QY 3095 CATTCGTGCATGCATGTTAGTCGGAAGTCCCGGGGTCAATATGTCCAAATGGTCTT 3154
DB CATTCGTGCATGCATGTTAGTCGGAAGTCCCGGGGTCAATATGTCCAAATGGCCTT 3000
QY 3155 CATGAAGCTGGCGGCTGACAGGTACGTACGTTTATAACCATCTTACCCCACTCGGGGA 3214
DB CATGAAGCTGGCGGCTGACAGGTACGTACGTTTATAACCATCTTACCCCACTCGGGGA 3060
QY 3215 CTGGGCCACGCGGCTACGAGACTTTCGGTGGGTAGAGCCGCTCTCTCTCGC 3274
DB CTGGGCCACGCGGCTACGAGACTTTCGGTGGGTAGAGCCGCTCTCTCTCGC 3120
QY 3275 CATGAGACCAAGGTATCATCCTGGGGAGCAGACCGCTGCTGTGGGACATCATCTT 3334
DB CATGAGACCAAGGTATCATCCTGGGGAGCAGACCGCTGCTGTGGGACATCATCTT 3180
QY 3335 GGGTCTACCCGCTCTCGGCCGAAGGGGGAAGAGATATTTTGGACCGGCTGATGTCT 3394
DB GGGTCTACCCGCTCTCGGCCGAAGGGGGAAGAGATATTTTGGACCGGCTGATGTCT 3240
QY 3395 CGAAGGCAAGGTGGCGACTCTTTCGCCCATACGCGCTACTCCCAACAAACGGGG 3454
DB CGAAGGCAAGGTGGCGACTCTTTCGCCCATACGCGCTACTCCCAACAAACGGGG 3300
QY 3455 CGTACTTGGTTCATCATCTAGCTACAGGCGGGACAAGAACCAAGGTCCGAAGGGA 3514
DB CGTACTTGGTTCATCATCTAGCTACAGGCGGGACAAGAACCAAGGTCCGAAGGGA 3360
QY 3515 GGTTCAGGTGGTTTCTACCGCAACAAATCTTTCTTCTGGCGACTGATCAACCGGCTGTG 3574
DB GGTTCAGGTGGTTTCTACCGCAACAAATCTTTCTTCTGGCGACTGATCAACCGGCTGTG 3420

QY 3575 CTGACTGTCTACCATGGCGCTGGCTCGAAGACCCCTAGCGGTCAAAAGGTCCAATCAC 3634
DB CTGACTGTCTACCATGGCGCTGGCTCAAGAGACCCCTAGCGGTCAAAAGGTCCAATCAC 3480
QY 3635 CAAAATGTACCAATGTAGACCTTGGACCTTGTGGCTGGCAGGCGCCCCCGGGGGCG 3694
DB CAAAATGTACCAATGTAGACCTTGGACCTTGTGGCTGGCAGGCGCCCCCGGGGGCG 3540
QY 3695 CTCATCACACCATGACAGCTGTGACCTCGGACCTTTACTTTGTCACGAGACATGCTGA 3754
DB CTCATCACACCATGACAGCTGTGACCTCGGACCTTTACTTTGTCACGAGACATGCTGA 3600
QY 3755 TGTCAATTCGCGTGGCGCGGAGCGCAGCAGGAGGGAAGTCTACTCTCCCGCAGGCCCGT 3814
DB TGTCAATTCGCGTGGCGCGGAGCGCAGCAGGAGGGAAGTCTACTCTCCCGCAGGCCCGT 3660
QY 3815 CTCCTACTGAAAGGCTCTCTCGGTGTGCATTTGCTTTGCCCTTCGGGGGACAGCTGTGG 3874
DB CTCCTACTGAAAGGCTCTCTCGGTGTGCATTTGCTTTTGGCCCTTCGGGGGACAGCTGTGG 3720
QY 3875 GGTCTTCGGGCTGTCTGTGTGCACCCCGGGGGTCCGGAAGCGGTGGACTTTCATACCCGT 3934
DB GGTCTTCGGGCTGTCTGTGTGCACCCCGGGGGTTCGGAAGCGGTGGACTTTCATACCCGT 3780
QY 3935 TGAGTCTATGGAATACTACCATGCGGTCTCGGTCTTTCACAGACAACTCAACCCCGCGC 3994
DB TGAGTCTATGGAATACTACCATGCGGTCTCGGTCTTTCACAGACAACTCAATCCCGCGC 3840
QY 3995 TGTAACCGAGACATTCOAAGTGGCACATCTGCAGCTCTTACTGGCAGCGCAAGAGCAC 4054
DB TGTAACCGAGACATTCOAAGTGGCACATCTGCAGCTCTTACTGGCAGCGCAAGAGCAC 3900
QY 4055 CAAAGTCGCGCTGCTGTATGCAGCCCAAGGTTACAAGTGTCTCTGTAACCGCTCGT 4114
DB CAAAGTCGCGCTGCTGTATGCAGCCCAAGGTTACAAGTGTCTCTGTAACCGCTCGT 3960
QY 4115 TGCCGCCACCTTAGGGTTTGGGGGTATATGTTCAAAGCACAACGGTATCGACCCCTAAAT 4174
DB TGCCGCCACCTTAGGGTTTGGGGGTATATGTTCAAAGCACAACGGTATCGACCCCTAAAT 4020
QY 4175 CAGAACTGGGTAAAGGACCAATTACCAGGGCGGTCCATTACGTACTCCACTATGGCAA 4234
DB CAGAACTGGGTAAAGGACCAATTACCAGGGCGGTCCATTACGTACTCCACTATGGCAA 4080
QY 4235 GTTCTTTCGCGACGCTGGCTGTCTGGGGGCGCTATGACATCATATATGTGTGAGTG 4294
DB GTTCTTTCGCGACGCTGGCTGTCTGGGGGCGCGCTATGACATCATATATGTGTGAGTG 4140
QY 4295 CCACCTCAACTGACTCGACTACCATCTTTGGGCATCGGCACAGTCTTGACCAAGCGGAGAC 4354
DB CCACCTCAACTGACTCGACTACCATCTTTGGGCATTTGGGCACAGTCTTGACCAAGCGGAGAC 4200
QY 4355 GGCTGGAGCGCGCTGCTGTGCTCGCACCGCTACACCTCCGGGATCGGTACCGTGCC 4414
DB GGCTGGAGCGCGCTGTGTGCTGTGCGCACCGCTACACCTCCGGGATCGGTACCGTGCC 4260
QY 4415 ACACCCCAATACGAGGAATAGGCTGTCCAACTGGAGAGATCCCTTCTATGGCAA 4474
DB ACACCCCAATACGAGGAATAGGCTGTCCAACTGGAGAGATCCCTTCTATGGCAA 4320
QY 4475 AGCCATCCCAATGAGGCGATCAAGGGGGGAGGAGCTCTCATTTTCTGCCATTCGAAGAA 4534
DB AGCCATCCCAATGAGGCGATCAAGGGGGGAGGAGCTCTCATTTTCTGCCATTCGAAGAA 4380
QY 4535 GAAATGTGACAGCTCGCCGCAAGAGCTGACAGGCTCGGACTGAACCGTGTAGCATATTA 4594
DB GAAATGTGACAGCTCGCCGCAAGAGCTGACAGGCTCGGACTGAAACGCTGTAGCATATTA 4440
QY 4595 CCGGGGCTTGTATGTCCGTCATACCGCTTATCGGAGACGCTGTGTGTGGGCAACAGA 4654
DB CCGGGGCTTGTATGTCCGTCATACCGCTTATCGGAGACGCTGTGTGTGGGCAACAGA 4500
QY 4655 CGCTCAATGACGGGTTTCAACGGCGGATTTTGACTCAGTGATCGACTCAATACATGTGT 4714

QY 6875 CTCCACATTACAGCAGAGACGGCTAAGCTAGGCTGGCTAGAGGGTCTCCCCCTCTTT 6934
DB |||||
QY 6721 CTCCACATTACAGCAGAGACGGCTAAGCTAGGCTGGCTAGAGGGTCTCCCCCTCTTT 6780
DB |||||
QY 6935 AGCCAGCTCATCAGCTAGCAGTGTCTCGGCTCTTTTGAAGGCGACATGACATACCCA 6994
DB |||||
QY 6781 AGCCAGCTCATCAGCTAGCAGTGTCTCGGCTCTTTTGAAGGCGACATGACATACCCA 6840
DB |||||
QY 6995 CCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCTCTTGTGGCGGAGGAGATGGG 7054
DB |||||
QY 6841 CCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCTCTTGTGGCGGAGGAGATGGG 6900
DB |||||
QY 7055 CGGAACATCACTCCGCTGGAGTCAGAGAAATAAGGTAGTAAATCTTGGACTCTTTGCAACC 7114
DB |||||
QY 6901 CGGAACATCACTCCGCTGGAGTCAGAGAAATAAGGTAGTAAATCTTGGACTCTTTGCAACC 6960
DB |||||
QY 7115 GCTTCACGGGAGGGGATGAGAGGAGATATCCGTCGCGGGAGATCTTCGGAANAATC 7174
DB |||||
QY 6961 GCTTCACGGGAGGGGATGAGAGGAGATATCCGTCGCGGGAGATCTTCGGAANAATC 7020
DB |||||
QY 7175 CAGGAAGTTCCTCAGCGTTGCCATATGCGCATATGGGCACGCCGGACTACATCTCCACTGCT 7234
DB |||||
QY 7021 CAGGAAGTTCCTCAGCGTTGCCATATGGGCACGCCGGACTACATCTCCACTGTT 7080
DB |||||
QY 7235 AGAGTCTTGAAGGACCCGGACTACGTCTCCGCTGGGTACACGGATGCCATTGCCACC 7294
DB |||||
QY 7081 AGAGTCTTGAAGGACCCGGACTACGTCTCCGCTGGGTACACGGATGCCATTGCCACC 7140
DB |||||
QY 7295 TACCAAGGCTCTCCAATACCACTCCAGCGGAGAAAGGAGCGTGTCTTGACAGAATC 7354
DB |||||
QY 7141 TACCAAGGCTCTCCAATACCACTCCAGCGGAGAAAGGAGCGTGTCTTGACAGAATC 7200
DB |||||
QY 7355 CAATGTCTTCTGCTTGGCGGAGCTGCGCACTAAGACCTTCGGTAGCTCCGATCGTC 7414
DB |||||
QY 7201 CAATGTCTTCTGCTTGGCGGAGCTGCGCACTAAGACCTTCGGTAGCTCCGATCGTC 7260
DB |||||
QY 7415 GSCGCTGTATAGCGACGCGCACCGCCTCTTCCGCTGGCTCCGACGCGGTGACAA 7474
DB |||||
QY 7261 GSCGCTGTATAGCGACGCGCACCGCCTCTTCCGCTGGCTCCGACGCGGTGACAA 7320
DB |||||
QY 7475 AGGATCCGACGTTGAGTCTACTCTCTCATGCCCCCTTGAAGGGAGCGCGGGGACCC 7534
DB |||||
QY 7321 AGGATCCGACGTTGAGTCTACTCTCTCATGCCCCCTTGAAGGGAGCGCGGGGACCC 7380
DB |||||
QY 7535 CGATCTCAGCAGCGGTCTTGTCTACCGTAGTGAGGAGGCTAGTGAGGATGTCGTCG 7594
DB |||||
QY 7381 CGATCTCAGCAGCGGTCTTGTCTACCGTAGTGAGGAGGCTAGTGAGGACGTCGTCG 7440
DB |||||
QY 7595 CTGCTCAATGTCCTATACGCTGACAGCGGCCCTGTATCACGCCATCGCTGCGGAGGAAAG 7654
DB |||||
QY 7441 CTGCTCAATGTCCTATACGCTGACAGCGGCCCTGTATCACGCCATCGCTGCGGAGGAAAG 7500
DB |||||
QY 7655 TAAGCTGCCCATCAACCCGTTGAGCAACTCTTTGCTGCTGCTACCAACAATGGTCTACGC 7714
DB |||||
QY 7501 TAAGCTGCCCATCAACCCGTTGAGCAACTCTTTGCTGCTGCTACCAACAATGGTCTACGC 7560
DB |||||
QY 7715 CACAACATCCCGAGGCGGAAGCCTCCGGCAGAGAAGGTCACTTTGACAGATTCGAAGT 7774
DB |||||
QY 7561 CACAACATCCCGAGGCGGAAGCCTCCGGCAGAGAAGGTCACTTTGACAGATTCGAAGT 7620
DB |||||
QY 7775 CTTGGATGATCAATTACCGGAGCTACTCAAGGAGATGAAGGCGMAGGGCTGCCAGTTAA 7834
DB |||||
QY 7621 CTTGGAGACCAATTACCGGAGCTGTCAAGGAGATGAAGGCGMAGGGCTGCCAGTTAA 7680
DB |||||
QY 7835 GGCTAAGCTTTCTATCTATAGAGAGGCGCTGCAAGCTGACGCCGCCACATTCGGGCCAAATC 7894
DB |||||
QY 7681 GGCTAAGCTTTCTATCTATAGAGAGGCGCTGCAAGCTGACGCCGCCACATTCGGGCCAAATC 7740
DB |||||
QY 7895 CAAATTTGGCTATGGGCGAAAGGACGTCGCGGAACCTTATCAGCAGGCGCGTTAAACCAT 7954
DB |||||
QY 7741 CAAATTTGGCTATGGGCGAAAGGACGTCGCGGAACCTTATCAGCAGGCGCGTTAAACCAT 7800
DB |||||

RESULT 15

DQ071885

LOCUS

DEFINITION

Hepatitis C virus polyprotein mRNA, complete cds.
9605 bp mRNA linear
VRL 20-JUN-2005

QY 7955 CCGCTCCGTTGGGAGGACTTGTCTGGAAGACACTGAAACCAATTCACACCAATCAT 8014
DB |||||
QY 7801 CCGCTCCGTTGGGAGGACTTGTCTGGAAGACACTGAAACCAATTCACACCAATCAT 7860
DB |||||
QY 8015 GGCAAAAGTGAAGTCTTCTCGCTCCAAACAGAGAGGGAGGCCGCAAGCCAGCTCGCT 8074
DB |||||
QY 7861 GGCAAAAGTGAAGTCTTCTCGCTCCAAACAGAGAGGGAGGCCGCAAGCCAGCTCGCT 7920
DB |||||
QY 8075 TATCGTATTTCCAGACCTCGGAGTTCGTGTATGGAGAAGATGCCCTTTTACGACGTGGT 8134
DB |||||
QY 7921 TATCGTATTTCCAGACCTCGGAGTTCGTGTATGGAGAAGATGCCCTTTTACGACGTGGT 7980
DB |||||
QY 8135 CTCACACCTTCTCAGCGCGTGATGGCTCTCTATACCGATTTCAATATCTCCCCAAGCA 8194
DB |||||
QY 7981 CTCACACCTTCTCAGCGCGTGATGGCTCTCTATACCGATTTCAATATCTCCCCAAGCA 8040
DB |||||
QY 8195 GCGGTCGAGTTCCTGCTGAATACCTGGAATCAAAGAAATGCCCTATGGCTTCTCAT 8254
DB |||||
QY 8041 GCGGTCGAGTTCCTGCTGAATACCTGGAATCAAAGAAATGCCCTATGGCTTCTCAT 8100
DB |||||
QY 8255 TGACACCGCTGTTTTCACCTCAACGCTCACTGAGAGTGACATTCGTGTGAGGAGTCAAT 8314
DB |||||
QY 8101 TGACACCTGCTGTTTTCACCTCAACGCTCAACGAGAGTGAATTCGTGTGAGGAGTCAAT 8160
DB |||||
QY 8315 TTACCAATGTTGTGACTTGGCCCCCGAGGCCAGACAGCCATAAGTGGCTCACAGACG 8374
DB |||||
QY 8161 TTACCAATGTTGTGACTTGGCCCCCGAGGCCAGACAGCCATAAGTGGCTCACAGACG 8220
DB |||||
QY 8375 GCTTTACATCGGGGTCCTTGACTTAATACTCAAAGGGCAGAACTGCGGTTATCGCGGTG 8434
DB |||||
QY 8221 GCTTTACATCGGGGTCCTTGACTTAATACTCAAAGGGCAGAACTGCGGTTATCGCGGTG 8280
DB |||||
QY 8435 CCGCGAAGTGGCTGCTGACGACTAGCTGCGGTAAATACCTCATCTGTTTGAAGGC 8494
DB |||||
QY 8281 CCGCGAAGTGGCTGCTGACGACTAGCTGCGGTAAATACCTCATCTGTTTGAAGGC 8340
DB |||||
QY 8495 CACTGACGCTGTGACGCTGCAAGCTCCAGGACTGCAAGTGTCTGTAACGAGACGA 8554
DB |||||
QY 8341 CACTGCGGCTGTGACGCTGCAAGCTCCAGGACTGCAAGTGTCTGTAACGAGACGA 8400
DB |||||
QY 8555 CTTTGTCTGTTATCTGTGAAGCGCGGGAACCCAGAGGATCGGGGCGCTTACGAGCCTT 8614
DB |||||
QY 8401 CTTTGTCTGTTATCTGTGAAGCGCGGGAACCCAGAGGATCGGGGCGCTTACGAGCCTT 8460
DB |||||
QY 8615 CACGAGGCTATGACTAGGTATTCGCCCCCGCGGGATCCGCCCAACAGAAATACGA 8674
DB |||||
QY 8461 CACGAGGCTATGACTAGGTATTCGCCCCCGCGGGATCCGCCCAACAGAAATACGA 8520
DB |||||
QY 8675 CTTGAGCTGATAAATCATATGTTCTCCAATGTGTGAGTCGCGCACGATGATCTGGCAA 8734
DB |||||
QY 8521 CTTGAGCTGATAAATCATATGTTCTCCAATGTGTGAGTCGCGCACGATGATCTGGCAA 8580
DB |||||
QY 8735 AAGGTATACCTACCTGACCCCGTGAACCCACACCCCTTGGACGGGCTGGTGGGAGAC 8794
DB |||||
QY 8581 AAGGTATACCTACCTGACCCCGTGAACCCACACCCCTTGGACGGGCTGGTGGGAGAC 8640
DB |||||
QY 8795 AGCTAGACACACTCCAATCAACTCTTGGCTAGGCAATATCATGATGCGCCACCCCT 8854
DB |||||
QY 8641 AGCTAGACACACTCCAATCAACTCTTGGCTAGGCAATATCATGATGCGCCACCCCT 8700
DB |||||
QY 8855 ATGGCAAGGATGATTCGTGATGACTCACTTTTCTCCAATCTTCTAGCTCAAGAGCACT 8914
DB |||||
QY 8701 ATGGCAAGGATGATTCGTGATGACTCACTTTTCTCCAATCTTCTAGCTCAAGAGCACT 8760
DB |||||
QY 8915 TGAAGAACCCCTGGATTCTCA 8935
DB |||||
QY 8761 TGAAGAACCCCTGGATTCTCA 8781
DB |||||

Db 841 GGAAC TTGCGCGGTGCTCTCTTTTATCTCTTCCCTTGGCTCTGCTGTCCTGTTTGACCA 900
QY 901 TCCAGACTTCGCTTATGAAGTGGCGCAACGCTGTCGGGATATACCATGTGCAGAACGACT 960
Db 901 TCCAGACTTCGCTTATGAAGTGGCGCAACGCTGTCGGGATATACCATGTGCAGAACGACT 960
QY 961 GCTCCAACTCAAGCATTTGTATGAGGACGCGGACGCTGATCATGCATACCTCCCGGCTGCG 1020
Db 961 GCTCCAACTCAAGCATTTGTATGAGGACGCGGACGCTGATCATGCATACCTCCCGGCTGCG 1020
QY 1021 TGCCCTGTGTTTCAGAGGGTAAACAGCTCCCTTGTGCTGGTAGCGCTCACTCCCAAGCTCG 1080
Db 1021 TGCCCTGTGTTTCAGAGGGTAAACAGCTCCCTTGTGCTGGTAGCGCTCACTCCCAAGCTCG 1080
QY 1081 CGGCGAGGAATGCCAGCGTCCCACTACGACAAATACGACGCCAGCTCGACTTGTCTGTTG 1140
Db 1081 CGGCGAGGAATGCCAGCGTCCCACTACGACAAATACGACGCCAGCTCGACTTGTCTGTTG 1140
QY 1141 GGAGGCTGTCTTGTCTCCGCTATGTACGTGGGGATCTCTGGGATCTATTTTCCCTCG 1200
Db 1141 GGAGGCTGTCTTGTCTCCGCTATGTACGTGGGGATCTCTGGGATCTATTTTCCCTCG 1200
QY 1201 TCTCCAGCTGTTCACCTTCTCGCTCGCGGCATGAGACAGTGCAGGACTGCACACTGCT 1260
Db 1201 TCTCCAGCTGTTCACCTTCTCGCTCGCGGCATGAGACAGTGCAGGACTGCACACTGCT 1260
QY 1261 CAATCTATCCCGCCATGTATCAGGTTCACCGATGGCTTGGGATATGATGATGACTGGT 1320
Db 1261 CAATCTATCCCGCCATGTATCAGGTTCACCGATGGCTTGGGATATGATGATGACTGGT 1320
QY 1321 CACCTTACAAACAGCCTAGTGTGTGCGAGTTGCTCCGGATCCCAAGCTGCTGTGGACA 1380
Db 1321 CACCTTACAAACAGCCTAGTGTGTGCGAGTTGCTCCGGATCCCAAGCTGCTGTGGACA 1380
QY 1381 TGGTGGCGGGGCCCACTGGGAGTCTCTGGGGGCTTGGCTACTATTTCCATGTTAGGGA 1440
Db 1381 TGGTGGCGGGGCCCACTGGGAGTCTCTGGGGGCTTGGCTACTATTTCCATGTTAGGGA 1440
QY 1441 ACTGGGCTAAGGTTCTGATTTGGGCGCTACTTGTCCGGGTGACGGGAGACCCACA 1500
Db 1441 ACTGGGCTAAGGTTCTGATTTGGATGCTACTTGTCCCGGGCTTTACGTCCTCTTTAATCTGGGT 1560
QY 1501 CGACGGGAGGTTGGCGGCACACACCTCCGGGTTACAGTCCCTTTTCTCATCTGGGG 1560
Db 1501 TGACAGGGGGACGATGCGCAAAACACCCCGGCTTTACGTCCTCTTTAATCTGGGT 1560
QY 1561 CGTCTCAGAAAATCCAGCTTTGGAATACCAACGGCAGCTGGCACATCAACAGGACTGCC 1620
Db 1561 CATCCAGAACATTCAGCTTTGTAACATCAACGGCAGCTGGCACATCAACAGGACAGCCC 1620
QY 1621 TAAATTTGCAATGACTCCCTCCAACTGGGTTCTTGGCGCGCTGTTTACGCACACAAGT 1680
Db 1621 TGAATGCAACAGATTCCTCTCAACACGGGTTTATTTGCTGCGGTGTTTACACGCACAAGT 1680
QY 1681 TCAACTCTGATCTGGATGCCAGAGCGATGGCCACGCTGGCCGCCATCGACAAGTTTCGATC 1740
Db 1681 TCAACTCTGATCTGGATGCCAGAGCGATGGCCACGCTGGCCGCCATCGACAAGTTTCGATC 1740
QY 1741 AGGGGTGGGGCCCCCATCACTTACTAAGACCTTAAACAGCTTCGAGTCAGAGGCTTATTGCT 1800
Db 1741 AGGGGTGGGGCCCCCATCACTTACTAAGTCTCAGAGCTCGGATCAGAGGCTTATTGCT 1800
QY 1801 GGCATTACGCGCTCGACCGTGTGTGCTGATACCGCGTFCGAGGTGTGTGGTCCAGTGT 1860
Db 1801 GGCATTACGCGCTCGACCGTGTGTGCTGATACCGCGTFCGAGGTGTGTGGTCCAGTGT 1860
QY 1861 ATTGTTTCAACCCAGCCCTGTGTGGTGGGACCAACCGATCGTTCCGCTACGTT 1920
Db 1861 ATTGTTTCAACCCAGCCCTGTGTGGTGGGACCAACCGATCGTTCCGCTACGTT 1920
QY 1921 ATAGCTGGGGGAGGAATGAGACAGAGCTGATGCTCTCTCAACAAACAGCGTCCGCGCACAA 1980
Db 1921 ATAGCTGGGGGAGGAATGAGACAGAGCTGATGCTCTCTCAACAAACAGCGTCCGCGCACAA 1980

QY 1981 GCAACTGGTTCGGCTGTACATGGATGAATAGTACTTGGGTTCACTAAGACGTCGGAGGTC 2040
Db 1981 GCAACTGGTTCGGCTGTACATGGATGAATAGTACTTGGGTTCACTAAGACGTCGGAGGTC 2040
QY 2041 CCCCGTGTAAACATCGGGGGGTTCGGTAAACCGCAACCTTGTGATCTGCCCCACGACTGCTTC 2100
Db 2041 CCCCGTGTAAACATCGGGGGGTTCGGTAAACCGCAACCTTGTGATCTGCCCCACGACTGCTTC 2100
QY 2101 GGAAGCACCCGAGGCTTACTTACAAAAATGTGGCTCGGGGCCATGTTGTACACCTAGGT 2160
Db 2101 GGAAGCACCCGAGGCTTACTTACAAAAATGTGGCTCGGGGCCATGTTGTACACCTAGGT 2160
QY 2161 GCCTAGTAGACTACCCCATACAGCTTTGGCACTACCCCTGCACCTCTCAATTTTTTCCATCT 2220
Db 2161 GCATAGTAGACTACCCCATACAGCTTTGGCACTACCCCTGCACCTCTCAATTTTTTCCATCT 2220
QY 2221 TTAAGGTTAGTAGTATGTGGGGGGTGGAGCACAGGCTCAATGCGCATGCAATTTGGA 2280
Db 2221 TTAAGGTTAGTAGTATGTGGGGGGTGGAGCACAGGCTCAATGCGCATGCAATTTGGA 2280
QY 2281 CTCGAGGAGCGCTGTAACTTGGAGACAGGATAGGTGAGAACTCAGCCCCGCTGCTGC 2340
Db 2281 CTCGAGGAGCGCTGTAACTTGGAGACAGGATAGGTGAGAACTCAGCCCCGCTGCTGC 2340
QY 2341 TGTCTACAAACAGAGTGGCAGATCTGCCCTGTCTTTCACACCCCTACCGGCTTTATCCA 2400
Db 2341 TGTCTACAAACAGAGTGGCAGATCTGCCCTGTCTTTCACACCCCTACCGGCTTTATCCA 2400
QY 2401 CTGCTTTGATCCATCTCCATCAGAACATCTGGACGTGCAATACCTGTACGGTGTAGGT 2460
Db 2401 CTGCTTTGATCCATCTCCATCAGAACATCTGGACGTGCAATACCTGTACGGTGTAGGT 2460
QY 2461 CAGGTTTGTCTCTTTTGCAATCAATGGGAGTACATCTCTGTGCTTTTCTCTCTCTGG 2520
Db 2461 CAGGTTTGTCTCTCTTTTGCAATCAATGGGAGTACATCTCTGTGCTTTTCTCTCTCTGG 2520
QY 2521 CAGACGCGCGTGTGCTGCTTGTGGATGATGCTGCTGATAGCCAGGCTCAGGCGCG 2580
Db 2521 CAGACGCGCGTGTGCTGCTTGTGGATGATGCTGCTGATAGCCAGGCTCAGGCGCG 2580
QY 2581 CTTTAGAGAACTTGGTGTCTCTCAATCGCGGCTCCGTGGCGGAGCGCATGTTCTCT 2640
Db 2581 CTTTAGAGAACTTGGTGTCTCTCAATCGCGGCTCCGTGGCGGAGCGCATGTTCTCTCT 2640
QY 2641 CTTTCTTGTGTTCTTCTGCGCCCTGGTACATTAAGGAGAGGCTGCTCTCTGGGGCGG 2700
Db 2641 CTTTCTTGTGTTCTTCTGCGCCCTGGTACATTAAGGAGGAGGCTGCTCTCTGGGGCGG 2700
QY 2701 CGTATGCTTTTATGGCGTATGGCGCTGCTCTGCTCTCTACTGGCGTTACACACGAG 2760
Db 2701 CGTATGCTTTTATGGCGTATGGCGCTGCTCTGCTCTCTACTGGCGTTACACACGAG 2760
QY 2761 CTTACGCTTTGGACCGGAGATGCTGCATCGTGGGGGTGCGGTTCTGTAGGCTCTGG 2820
Db 2761 CTTACGCTTTGGACCGGAGATGCTGCATCGTGGGGGTGCGGTTCTGTAGGCTCTGG 2820
QY 2821 TATTCTTGACCTTGTACACATPACAAAAGTGTTCATCTAGGCTCATATGTGGTGTAC 2880
Db 2821 TACTCTTGACCTTGTACACATPACAAAAGTGTTCCTCGCTAGGCTCATATGTGGTGTAC 2880
QY 2881 AATACTTTATCAGAGCGCGAGGACATGCAAGTGTGGTCCCGCTCCCGCTCAACGTTTC 2940
Db 2881 AATATTTTATCAGAGCGCGAGGACATGTCGAAGTGTGGATCCCGCTCCCGCTCAACGTTTC 2940
QY 2941 GGGGAGGCGCGGATGCCATCATCTCTCCTCAGCTGCGGTTTCATCCAGAGTTAATTTTG 3000
Db 2941 GGGGAGGCGCGGATGCCATCATCTCTCCTCAGCTGCGGATCCATCCAGAGTTAATTTTG 3000
QY 3001 ACATCACAAAACCTCTGCTCGCCATATCTCGGCGCGCTCATGTGCTCCAGGCTGCATAA 3060
Db 3001 ACATCACAAAACCTCTGCTCGCCATATCTCGGCGCGCTCATGTGCTCCAGGCTGCATAA 3060

QY 3061 CGAGAGTCCGCTACTTTCGTCGCGCTCAAGGGCTCAATTCGTGATGCAATGTTAGTCCGAA 3120
 Db |||||
 QY 3061 CGAGAGTCCGCTACTTTCGTCGCGCTCAAGGGCTCAATTCGTGATGCAATGTTAGTCCGAA 3120
 Db |||||
 QY 3121 AAGTCCCGGGGTCAATTAATGTCCTCAATGCTTTCATGAAGTGGCGCGCTGACAGGTA 3180
 Db |||||
 QY 3121 AAGTCCCGGGGTCAATTAATGTCCTCAATGCTTTCATGAAGTGGCGCGCTGACAGGTA 3180
 Db |||||
 QY 3181 CGTACGTTTATAACCACTTTACCCCACTGGGGACTGGGCCCAACGCGGCTTACGAGACC 3240
 Db |||||
 QY 3181 CGTACGTTTATAACCACTTTACCCCACTGGGGACTGGGCCCAACGCGGCTTACGAGACC 3240
 Db |||||
 QY 3241 TTGCGGTGGGTTAGAGCCCGTCTTCTTCGCCCATGAGACCAAGTCACTACCTGGG 3300
 Db |||||
 QY 3241 TTGCGGTGGGTTAGAGCCCGTCTTCTTCGCCCATGAGACCAAGTCACTACCTGGG 3300
 Db |||||
 QY 3301 GAGCAGACCCGTCGCTGGTGGGACATCATCTTGGGTCTACCCGCTTCGCCCGCAAGGG 3360
 Db |||||
 QY 3301 GAGCAGACCCGTCGCTGGTGGGACATCATCTTGGGTCTACCCGCTTCGCCCGCAAGGG 3360
 Db |||||
 QY 3361 GGAAGGAGATATTTTGGGACCCGCTGATAGTCTCGAAGGGCAAGGGTGGCGACTCCTTTG 3420
 Db |||||
 QY 3361 GGAAGGAGATATTTTGGGACCCGCTGATAGTCTCGAAGGGCAAGGGTGGCGACTCCTTTG 3420
 Db |||||
 QY 3421 CGCCCATACGCGCTACTTCCCAACAAACGCGGGCGTACTTGGTTGCAATCATCACTAGCC 3480
 Db |||||
 QY 3421 CGCCCATACGCGCTACTTCCCAACAAACGCGGGCGTACTTGGTTGCAATCATCACTAGCC 3480
 Db |||||
 QY 3481 TCACAGCCGGGCAAGAACCAAGTTCGAGGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
 Db |||||
 QY 3481 TCACAGCCGGGCAAGAACCAAGTTCGAGGGGAGGTTCAAGTGGTTTCTACCGCAACAC 3540
 Db |||||
 QY 3541 AATCTTTCTCGGCACTGATCAACGCGCTGTCTGGAAGTCTACCATGGGCGCTGGCT 3600
 Db |||||
 QY 3541 AATCTTTCTCGGCACTGATCAACGCGCTGTCTGGAAGTCTACCATGGGCGCTGGCT 3600
 Db |||||
 QY 3601 CGAAGACCTTAGCGGTCCAAAAGGTCCAAATCAACCCAAATGTACACCAATGTAGACTGG 3660
 Db |||||
 QY 3661 ACCTGTCGCTGGCAGCGCCCGCGGGCGCTCCATGACACCAATGAGCTGTGGCA 3720
 Db |||||
 QY 3661 ACCTGTCGCTGGCAGCGCCCGCGGGCGCTCCATGACACCAATGAGCTGTGGCA 3720
 Db |||||
 QY 3721 GCTCGGACCTTACTTCTGTCACGAGACATGCTGATGTCATTTCGCTGCGCGCGAGGCG 3780
 Db |||||
 QY 3721 GCTCGGACCTTACTTCTGTCACGAGACATGCTGATGTCATTTCGCTGCGCGCGAGGCG 3780
 Db |||||
 QY 3781 ACAGCAGGGGAAGTCTACTCTCCCCAGGCGCGTCTCTACCTGAAAGGCTCTCTCGGGTG 3840
 Db |||||
 QY 3781 ACAGCAGGGGAAGTCTACTCTCCCCAGGCGCGTCTCTACCTGAAAGGCTCTCTCGGGTG 3840
 Db |||||
 QY 3841 GTCCATTGCTTTGCGGTCGAGTCTTACCCCGTGGGCGCTTCCGCGGCTGCTGTGACACC 3900
 Db |||||
 QY 3841 GTCCATTGCTTTGCGGTCGAGTCTTACCCCGTGGGCGCTTCCGCGGCTGCTGTGACACC 3900
 Db |||||
 QY 3901 GGGGGTTCGAGAGCGGTGACTTACACCCGTTGAGTCTATCGAAACTACCATTCGCGT 3960
 Db |||||
 QY 3901 GGGGGTTCGAGAGCGGTGACTTACACCCGTTGAGTCTATCGAAACTACCATTCGCGT 3960
 Db |||||
 QY 3961 CTCGCGTCTTTCACAGACAACTCAACCCCGCGGTGTACCCGACAGACATTCGAAAGTGGCAC 4020
 Db |||||
 QY 3961 CTCGCGTCTTTCACAGACAACTCAACCCCGCGGTGTACCCGACAGACATTCGAAAGTGGCAC 4020
 Db |||||
 QY 4021 ATCTGCAAGTCTTCTGAGAGCGGCAAGACCAAGAGTCCGCGCTGCGTATGAGCC 4080
 Db |||||
 QY 4021 ATCTGCAAGTCTTCTGAGAGCGGCAAGACCAAGAGTCCGCGCTGCGTATGAGCC 4080
 Db |||||
 QY 4081 AAGGGTCAAGGTGCTCTGTCGAGCCGCTGCTGCGCCACCTTATAGGTTTGGGGCGT 4140
 Db |||||
 QY 4081 AAGGGTCAAGGTGCTCTGTCGAGCCGCTGCTGCGCCACCTTATAGGTTTGGGGCGT 4140
 Db |||||
 QY 4141 ATATGTCCAGGCAACGCTATCGAACCCTTAACATCAGAACTGGGGTAAGGACCATTACCA 4200
 Db |||||

4141 ATATGTCCAGGCAACAGGTATCGACCTTAACATCAGAACTGGGGTAAGACCATTACCA 4200
 QY |||||
 4201 CGGGCGGCTCCATTACGTAATCCACTATGGCAAGTTCCTTCCGACGCTGCTGTCTTG 4260
 Db |||||
 4201 CGGGCGGCTCCATTACGTAATCCACTATGGCAAGTTCCTTCCGACGCTGCTGTCTTG 4260
 Db |||||
 4261 GGGGCGCTTACGATCAATATATGTGATGAGTGGCCACTCAACTGACTCGACTTACCATCT 4320
 QY |||||
 4261 GGGGCGCTTACGATCAATATATGTGATGAGTGGCCACTCAACTGACTCGACTTACCATCT 4320
 Db |||||
 4321 TGGGATCGGACAGCTCTGGACCAAGCGGAGACGCTGGAGCGGCTCGTCTGCTCG 4380
 QY |||||
 4321 TGGGATCGGACAGCTCTGGACCAAGCGGAGACGCTGGAGCGGCTCGTCTGCTCG 4380
 Db |||||
 4381 CCAACGCTACCTCCGCGATCCGTTACCGTGGCAACCCCAATATCGAGGAATAGGCC 4440
 QY |||||
 4381 CCAACGCTACCTCCGCGATCCGTTACCGTGGCAACCCCAATATCGAGGAATAGGCC 4440
 Db |||||
 4441 TGTCCCAATCGAGAGATCCCTTCTATGGCAAGCCCATCCCATTTGAGGCCCATCAAGG 4500
 QY |||||
 4441 TGTCCCAATCGAGAGATCCCTTCTATGGCAAGCCCATCCCATTTGAGGCCCATCAAGG 4500
 Db |||||
 4501 GGGGAGGCACTCTCATTTCTGCCATTCCAAGAAAGAAATGTGACGAGCTCGCCGCAAGC 4560
 QY |||||
 4501 GGGGAGGCACTCTCATTTCTGCCATTCCAAGAAAGAAATGTGATGAGCTCGCCGCAAGC 4560
 Db |||||
 4561 TGAACGCTCGGACTGAAACGCTGTAGCATATTTACCGGGGCTTGTATGCTCGTCTATAC 4620
 QY |||||
 4561 TGTCCGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTGTATGCTCGTCTATAC 4620
 Db |||||
 4621 CGCCTATCGGAGAGCTGCTTCTGTCGCAACAGACGCTCTAATGACGGGTTTCAACGGCG 4680
 QY |||||
 4621 CAACTACGAGAGAGCTCATTTGCTGAGCAACGACGCTCTAATGACGGGTTTCAACGGCG 4680
 Db |||||
 4681 ATTTGACTCAGTGATCGACTCAATACATGCTGTCCACAGACAGTGCAGCTTCAGCTTGG 4740
 QY |||||
 4681 ATTTGACTCAGTGATCGACTCAATACATGCTGTCCACAGACAGTGCAGCTTCAGCTTGG 4740
 Db |||||
 4741 ATCCCACTTTACCATTTGAGACGACACCGTGGCCCAAGACGCGGTTCGCGTTCGCAAC 4800
 QY |||||
 4741 ACCGACCTTACCATTTGAGACGACACCGTGGCCCAAGACGCGGTTCGCGTTCGCAAC 4800
 Db |||||
 4801 GGGAGGTAGAACTGGCAGGGGTAGAGTGGCATCTACAGGTTTGTGACTCCAGGAGAAC 4860
 QY |||||
 4801 GGGAGGTAGAACTGGTGGGAGGAGTGGGCTCTACAGGTTTGTGACTCCAGGAGAGC 4860
 Db |||||
 4861 GGGCTCGGCGATGTTGATTTCTCGTCTGCTGAGTGTATGACGCGGGCTGTGCTT 4920
 QY |||||
 4861 GGGCTCGGCGATGTTGATTTCTCGTCTGCTGAGTGTATGACGCGGGCTGTGCTT 4920
 Db |||||
 4921 GGTATGAGCTCAGCGCGCTGAGACCTCGGTTCGTTAGGTTTCGCGGCTTACCTAAATACACAG 4980
 QY |||||
 4921 GGTATGAGCTCAGCGCGCTGAGACCTCAGTTCGTTAGGTTTCGCGGCTTACCTAAATACACAG 4980
 Db |||||
 4981 GGTTCGCGCTGTCAGGACCACTCTGAGTTCGAGAGAGGCTTTCACAGGCTTCACCC 5040
 QY |||||
 4981 GGTTCGCGCTGTCAGGACCACTCTGAGTTCGAGAGAGGCTTTCACAGGCTTCACCC 5040
 Db |||||
 5041 ACATAGTGGCCTTCTCTGTCGAGACTAAACAGGCGAGAGACAACTTTCTTCTTACCTGG 5100
 QY |||||
 5041 ACATAGTGGCCTTCTCTGTCGAGACTAAACAGGCGAGAGACAACTTTCTTCTTACCTGG 5100
 Db |||||
 5101 TGGCATATCAAGCTTACAGTGTGGCCAGGGCTCAAGCTCCACCTCCATTCGTTGGGACCAA 5160
 QY |||||
 5101 TAGCATACAGGCTACGCTGTGGCCAGGGCTCAGGCTCCACCTCCATTCGTTGGGACCAA 5160
 Db |||||
 5161 TGTGAAAGTGTCTCATACGGCTGAAACCTACATGACGCGGCGCAACCCCTGTGTATA 5220
 QY |||||
 5161 TGTGAAAGTGTCTCATACGGCTGAAACCTACATGACGCGGCGCAACCCCTGTGTATA 5220
 Db |||||
 5221 GGCTAGGAGCGCTCCAAAATGAGGTCTATCTTCACACACCCCATTAACATACATCAGG 5280
 QY |||||

Db 5221 GGCTGGGAGCCGTCCAAAACGAGGTCACTACACACATCCCATAAACCAAAATACATCATGG 5280
Qy 5281 CATGATGTCGGCTGACCTGGAGGTGCTCACTAGCACCTGGTGTGCTAGCGGAGTCC 5340
Db 5281 CATGATGGCAGCTGACCTGGAGGTGCTACGAGCACCTGGTGTGCTAGCGGAGTCC 5340
Qy 5341 TTGACGCTTTGGCCGCTACTGCTGACGACGAGGCGAGTGTGGTCAATTTGGGCGAGGATCA 5400
Db 5341 TAGCAGCTCTGGCCGCTACTGCTGACAAAGGGCAGCGTGTCTATTTGGGCGAGGATCA 5400
Qy 5401 TCTTGTCCGGAGAGCCAGCTGCTGCTCCGACAGGAAAGTCTCTACAGAGGTTCGATG 5460
Db 5401 TCTTATCCGGAAAGCCGCTATCATTTCCGACAGGAAAGTCTCTACCGGGAGTTCGATG 5460
Qy 5461 AGATGGAAGAGTGTCCCTCAACAATTTCTTATCATCGACGAGGAAATGAGCTCGCCGAGC 5520
Db 5461 AAATGGAAAGAGTGCCTTCACACCTTCCTATCATCGAACAGGAAATGAGCTTCGCCGAAC 5520
Qy 5521 AATTCAAGCAAAAGGCGCTCGGTTTGTTCGAAACGGCCACCAAGCAAGCGGAGGCTGCTG 5580
Db 5521 AATTCAAGCAGAAGGCGCTCGGTTTGTTCGAAACAGCCACCAAGCAAGCGGAGGCTGCTG 5580
Qy 5581 CTCCTGTGTGAGTCCAAAGTGGCAGGCGCTTGAGACCTTCTGGGCGAAGCACATGTGGA 5640
Db 5581 CTCCTGTGTGAGTCCAAAGTGGCAGGCGCTTCGAAAGCCTTCTGGGCGAAGCATATGTGGA 5640
Qy 5641 ATTTTCATCAGCGGAATACAGTACCTAGCAGGCTTATCCACTCTGCTCGAAACCCCGGA 5700
Db 5641 ACTTCATCAGCGGATACAGTACTTAGCAGGCTTGTCCACTCTGCTGGGAACCCCGGA 5700
Qy 5701 TAGCATCATTTGATGGCATTTACAGCTTTCTATCACTAGCCCGCTCACCAACCCCAAAACACCC 5760
Db 5701 TAGCATCACTAATGGCATTTACAGCTCTATACACAGCCCGCTCACCAACCCCAAAACCC 5760
Qy 5761 TCTGTGTTAAACATCTTGGGGGGTGGTGGCGCGCCAACTTGTCCCGCCAGCGCTGCTT 5820
Db 5761 TCTGTGTTAAACATCTTGGGGGGTGGTGGCGCGCCAACTTGTCCCGCCAGCGCTGCTT 5820
Qy 5821 CAGCTTTCGTGGCGCGCATCGCGGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG 5880
Db 5821 CAGCTTTTGTGGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCCCTTGGGAAGG 5880
Qy 5881 TGCTCGTGACATCTTGGCGGCTATGGGCGAGGGTAGCCGCGCACTCTGTGGCCTTTA 5940
Db 5881 TGCTCGTGACATCTTGGCGGCTATGGGCGAGGGTAGCCGCGCACTCTGTGGCCTTTA 5940
Qy 5941 AGGTCAATAGCGCGAGGTGCCCTCACCGAGGACCTGTGTCAACTTATCTCCCTGCATCC 6000
Db 5941 AAATCATGAGCGCGAGATGCCCTCTCCGAGGACCTGTGTTAAACCTACTTCCCGCATCC 6000
Qy 6001 TCTCTCTGCTGCTGCTGCTGCTGGGTCGTGCGGACGCAATACTGTGCGGACGTGG 6060
Db 6001 TCTCTCTGCTGCTGCTGCTGCTGGGTCGTGCGGACGCAATACTGTGCGGACGTGG 6060
Qy 6061 GCCCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTCCGCGGGTA 6120
Db 6061 GCCCGGAGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTCCGCGGGTA 6120
Qy 6121 ACCAGTCTCCCTTACGCACTATGTGCTGAGAGGACGCTGAGCAGCTGTCACTCAGA 6180
Db 6121 ACCAGTCTCCCCACGCACTACGTGCTGAGAGGACGCTGAGAGGCTCAACCCAGA 6180
Qy 6181 TCTCTCTAGCCTTACCATCACTCACTGCTGAGCGGCTCCACGATGGATTAATGAGG 6240
Db 6181 TTCTCTCCAGCCTTACCATCACTCACTGCTGAGAGGCTCCACGATGGATTAATGAGG 6240
Qy 6241 ACTGCTTACGCCATGCTCGGCTCGTGGCTGAAGGATGTTTGGGATTTGGATATGCACGG 6300
Db 6241 ACTGCTTCCAGCCATGTTTCGGCTCGTGGCTCAGGGATGTTTGGGATTTGGATATGCACGG 6300
Qy 6301 TGTGATCACTTCAAGACTGCTGCTCCAGTCCAGTCCAAACTCTCTGCGCGGTTTACCGGAGTCC 6360
Db 6301 TGTGATCACTTCAAGACTGCTGCTCCAGTCCAGTCCAAAGCTCTCTGCTGCGGAGTTC 6360

Qy 6361 CTTTCTGTTCATGCGCAACGCGGGTACAAGGAGTCTTGGCGGGGAGCGGCATCATGCAAA 6420
Db 6361 CTTTCTTTTCATGCGCAACGTTGGATATAAGGGAGTCTGCGGGGAGACGGCATCATGCAAA 6420
Qy 6421 CCACCTGCCATGCGGAGCAAGATCGCCGGACATGTCAAAAACGGTTCATAGAGATCG 6480
Db 6421 CCACCTGCCATGCGGAGCAAAATCACCGGACATGTCAAAAACGGTTCATAGAGATCG 6480
Qy 6481 TAGGGCTTAGAACCTGACAGCAACGTTGGCAGCGAAGCTTCCCATCAACGATACACCA 6540
Db 6481 TTGGSCCTTAATAACCTGTAGCAACACGTGGCAGGAACATTTCCCTATCAACGATACACCA 6540
Qy 6541 CGGACCTTTGACACCTCTCCCGCGCCCAACTATTTCAGGGCGCTATGGCGGTGGCTG 6600
Db 6541 CAGCCCTTGACACCTCTCGCCAGCGCCAACTATTCCAAGCGTGTGGCGGTGGCGG 6600
Qy 6601 CTGAGGATGAGTGGAGGTTACGCGTGTGGGGATTTCCACTAGCTGACGGGATGACTA 6660
Db 6601 CCGAGGATGAGTGGAGGTTACGCGTGTGGGGATTTCCACTAGCTGACGGGATGACTA 6660
Qy 6661 CTGACACGTAAAGTGGCCATGCGGCTCGGCGCCCGAATTTCTTACGAGGTGGATG 6720
Db 6661 CTGACACATAAAATGCCCATGCGGCTCGGTCGGTCCCCGAGTCTTTCACGGAAATGGAGG 6720
Qy 6721 GAGTGGCTTGACACAGGTACGCTCCGGCTGCAAAACCTTCTTACGGGAGGACGTCAGT 6780
Db 6721 GGTGACGCTGACAGGTACGCCCCAGCATGCAAACTCTCTTACGGGATGAGGTACAT 6780
Qy 6781 TCCAGGTCGGGCTCAACAAATCTTGGTGGGTGCGAGTCCCATGGAGCGCCGAACCGG 6840
Db 6781 TCCAGGTCGGGCTCAACAGTACCGGTGGGTGCGAACTCCCATGTGAAACCGGAACCGG 6840
Qy 6841 ACCTAAAGTGTCTTACTTCCATGCTACCGATCCCTCCCATTTACAGCAGAGCGGCTA 6900
Db 6841 ACGTAGCAGTGTCTTACTTCCATGCTCACCGATCCACGACCCCTCCCATTTACAGCGAAACGGCTA 6900
Qy 6901 AGCGTAGGCTGGCTAGAGGTCTCCCCCTCTTTTAGCCAGCTCATCAGTACGCACTGTT 6960
Db 6901 AGCGTAGGCTGGCCAGGGATCTCCCCCTCTTTGGCCAGCTCATCAGTACGCACTGTT 6960
Qy 6961 CTGCGCTCTTTGAAAGCGACATGCACTACCCGCTACGACTCCCGGAGCGTGAACCTCA 7020
Db 6961 CTGCGCTCTCTTTGAAAGCGACATGCACTACCCGCTACGACTCCCGGAGCGTGAACCTCA 7020
Qy 7021 TCGAGGCCAACTCTTGTGGCGGAGAGTGGGGGAAACATCACTCCGTTGGAGTCA 7080
Db 7021 TCGAGGCCAACTCTTGTGGCGGAGAGTGGGGGAAACATCACTCCGCTGGAGTCA 7080
Qy 7081 AGAATAAGGTAGTAATTTCTGACTCTTTTCGAAACCGCTTTCAGCGGAGGGGATGAGGG 7140
Db 7081 AGAACAAGGTAGTAATTTTGGACTCTTTTGAACCACTTTCAGCGGAGGAGATGAGGG 7140
Qy 7141 AGATATCCGTCCGCGGAGATCTCGGAAATTCAGGAAATTCCTCAGCGTTGCCCA 7200
Db 7141 AAGTGTCCGTTCCGCGGAGATCTCGGAAATTCAGGAAATTCCTCAGCGTTGCCCA 7200
Qy 7201 TATGGGACCGCCGACTACAACTCTCCACTGTAGTCTTGGAGGACCGGACTAGG 7260
Db 7201 TATGGGACCGCCGAGTATCAATCCACCTTTTACAGCCCTTGAAGGACCGGACTAGG 7260
Qy 7261 TCCCTCCGTTGGTACACGGATGCCCATTCGCACTTACCAAGCTCTCTCAATACCACTTC 7320
Db 7261 TCCCTCCGTTAGTACACGGTGGCCATTTGCGGCTGTCTAAGGCCCTTCGATACCACTTC 7320
Qy 7321 CACGAGAAAGAGGACGTTGTCTGTGACAGAAATCAATGTGTCTTCTGCTTGGCGGAGC 7380
Db 7321 CACGAGGAGAGGACGTTGTCTGTGACAGTCTACCGTGTCTTCTGCTTGGCGGAGC 7380
Qy 7381 TCGCCACTAAGACCTTCCGTAGTCCGATCGTGGCGCTTGTATAGCGGACCGGAGTCC 7440
Db 7381 TCGCCAAAGACCTTTCGGAGCTCCGAATCGTGGCTGTGTGACAGCGGATGGCAACTG 7440

7441 CCCTTCTGACCTGGCTCCGACGACGGTGACAAAGGATCCGACGTTGAGTGGTACTCT 7500
7441 CCCTTCTGACGACCTCCGACGATGGAGATGACGATCCGACGTTGAGTGGTACTCT 7500
7501 CCATGCCCCCTTGAAGGAGCGCGGAGACCCGATCTCAGCAGCGGTCTTGTGCTA 7560
7501 CCATGCCCCCTTGAAGGAGCGCGGAGACCCGATCTCAGCAGCGGTCTTGTGCTA 7560
7561 CCGTAGTGAGGAGGTAGTGAGATGTCTGTCTGCTCAATGTCTTATACGTGACAG 7620
7561 CTGTGAGCGAGGAGGTAGTGAGAGCGTCTGTCTGTTTCGATGCTTACACATGACAG 7620
7621 GCGCCCTGATCAGCCATCGCTGCGGAGAAAGTAAAGTGCCTCCATCAACCCGTTGAGCA 7680
7621 GCGCCCTTAATCACACATGCGCGCGGAGAGACCAAGCTGCCATCAATGCACTGAGCA 7680
7681 ACTCTTTGCTCGCTACCAACAACATGCTCTACGCCACAACATCCCGCAGCGCAAGCTCC 7740
7681 ACTCTTTGCTCGCTACCAACAACATGCTCTATGCCACAACATCCCGAGCGCAAGCTGC 7740
7741 GGCAGAAAGAGTCACTTTTGACAGATTTGCAAGTCTCTGGATGATCAATACCGGAGCTAC 7800
7741 GGCAGAAAGAGTCACTTTTGACAGATCTCAGTCTCGGACGACCACTACCGGAGCTGC 7800
7801 TCAAGGAGATGAAGGCGAAGGCTCCACAGTTAAGCTAAGCTTCTATCTATAGAGAG 7860
7801 TCAAGGAGATGAAGGCGAAGGCTCCACAGTTAAGCTAAGCTATGCTATCCGTGGAGAG 7860
7861 CTGTCAAGCTGAGCCGCCACATTCGCGCAATCCAAATTTGGCTATGGGCGAAAGACG 7920
7861 CTGTGAAGCTGAGCCGCCACATTCGCGCGAGTCTAAATTTGGCTACGGGCGAAAGACG 7920
7921 TCCGGAACCTATCAGCAGGCGCGTTAAACACATCCGCTCCGTGGGAGGACTTGTGG 7980
7921 TCCGGAACCTATCAGCAAGCGCGTTAAACACATCCGCTCCGTGGGAGGACTTGTGG 7980
7981 AAGACTCTGAAACACCAATTTGACACCAACATCATATGCGCAAAATGAGGTCTTTGTGTC 8040
8041 AACAGAGAGGAGCGCGCAAGCAGCTCGCTTATCGTATCCAGAGCTGGGAGTTC 8100
8041 AACAGAGAGGAGGAGTTCGCAAGTCACTCGCTTATCGTATCCAGAGCTGGGAGTTC 8100
8101 GTGTATCCGAGAGATGCGCTTTACGACGTGGTCTCCACCTTCCTCAGCGCGTGATGG 8160
8101 GTGTGTCGAGAAATGCGCTTTACGATGTGCTTCAACCTCCCTCAGCGCGTGATGG 8160
8161 GCTCCTCATACGGATTTCAATACTCCCCCAAGCAGCGGTCGAGTTCCTGGTGAATACCT 8220
8161 GCTCCTCATACGGATTTCCAGTACTCTCTGCGGACGCGGTCGAGTTCCTGGTGAATACCT 8220
8221 GGAATCAAGAAATAGCCTTATGGCTTCTCATATGACACCGCGTGTGTTGACTCAACGG 8280
8221 GGAATCAAGAAAGCGCTATGGGCTTCGATATGACACCGCGTGTGTTGACTCAACGG 8280
8281 TCACTGAGTGACATTCGCTGTCAGAGTCAATTTTACCAATGTGACTTTGCGCCCGG 8340
8281 TCACTGAGTGACATTCGCTGTCAGAGTCAATTTTACCAATGTGACTTTGCGCCCGG 8340
8341 AGGCAGACAGCCATAGGTCTGCTCAGAGCGGCTTTTACATCGGGGTCCTCTGACTA 8400
8341 AAGCCAGACAGCCATAGGTCTGCTCAGAGCGGCTTTTACATCGGGGTCCTCTAACA 8400
8401 ACTCAAAAGGCGAAGTCCGGTATTCGCGGTGCGCGCAAGTGGCGTGTGAGACTA 8460
8401 ATTCAAAAGGCGAAGTCCGGCTATCCGCGTTCGCGCGAGCGGCGTGTGAGACTA 8460
8461 GCTGGGTAAACCTTCACTGTTTACTTGAAGGCCACTGACGCTGTGAGCTTCAAGC 8520
8461 GCTGGGTAAACCTTCACTGTTTACTTGAAGGCCCTGCGCGCTGTGAGCTTCAAGC 8520
8521 TCCAGGACTGACGATGCTCGTGAACGAGACGACCTTGTGCTTATCTGTGAAGCGCGG 8580

8521 TCAGGACTGACGATGCTCTGATGCGGGAACGACTTGTCTGTAAAGCGCGG 8580
8581 GAACCCAGGAGATGCGCGGCCCTACGAGCCTTACGAGGCTATGACTAGGTATTCG 8640
8581 GGACCCAGGAGACGAGGCGAACTCCGAGTCTTACGAGGCTATGACGAGTACTCTG 8640
8641 CCCCCCGGGGATCGCCCCCAACAGATACGACTGAGCTGATAACATCATGTTCT 8700
8641 CCCCCCGGGGATCGCCCCCAACAGATACGACTGAGCTGATAACATCATGTTCT 8700
8701 CCAATGTGTGAGTCCGCGCACGATGCTCTGGCAAAAGGGGTATACCTCACCCGTGAC 8760
8701 CAAATGTGTGAGTCCGCGCACGATGCTCTGGCAAAAGGGGTATACCTCACCCGTGAC 8760
8761 CCAACACCCCTTGTGACGCGCTGCGTGGGAGACGACTAGACACCTCCAATCACTCT 8820
8761 CCGCTCCCTTGTGCGCGGCTGCGTGGGAGACGACTAGACACCTCCAATCACTCT 8820
8821 GGTAGGCAATATCATGATGATGCGCCACCTATGGGCAAGGATGATTCGATGACTC 8880
8821 GGTAGGCAACATCATGATGATGCGCCACCTATGGGCAAGGATGATTCGATGACTC 8880
8881 ACTTTTCTCCATCTCTTAGCTCAAGAGCAACTTGAAGGAGGCTTGGATTTGTCAGATCT 8940
8881 ATTTCTTCTCCATCTCTTAGCTCAGGACAACTTGAAGGAGGCTTGGATTTGTCAGATCT 8940
8941 ACGGGCTTGTCTACTCCATTTGACCACTTGCACCTACTCAGATCATTTGAACGACTCCATG 9000
8941 ACGGGCTTGTCTACTCCATTTGACCACTTGCACCTACTCAGATCATTTGAACGACTCCATG 9000
9001 GTCTTAGCGATTTACACTCCAGTACTCTCCAGTGGATGATCAATAGGTTGGCTTCAT 9060
9001 GCCTTAGCGATTTTCACTCCATGATCTCTCCAGTGGATGATCAATAGGTTGGCTTCAT 9060
9061 GCCTTAGGAACTTGGGTACCACTTGCAGGCTGGAGACATCGGGCCAGAGTGTCC 9120
9061 GCCTTAGGAACTTGGGTACCACTTGCAGGCTGGAGACATCGGGCCAGAGTGTCC 9120
9121 GCGCTAAGCTACTGTCCCAGGGGGAGGCGCCACTTGTGGAGATACCTCTTTAACT 9180
9121 GCGCTAAGCTACTGTCCCAGGGGGAGGCGCCACTTGTGGAGATACCTCTTTAACT 9180
9181 GGGAGTAAAGGACAAAGCTTAAACTCACTCAATCCCGCGCGCTCCAGCTGGACTTGT 9240
9181 GGGAGTAAAGGACAAAGCTTAAACTCACTCAATCCCGCGCGCTCCAGCTGGACTTGT 9240
9241 CTGCTGTTGCTGCTGCTTCTGCTTCTCTCTACTCTCTGAGGGTAGGCAATTTACCTGC 9300
9241 CCAGCTGTTGCTGCTGCTTCTGCTTCTCTCTACTCTCTGAGGGTAGGCAATTTACCTGC 9300
9301 GACCCGCTGTTTCCGCTTCTGCTTCTCTCTACTCTCTGAGGGTAGGCAATTTACCTGC 9360
9301 GACCCGCTGTTTCCGCTTCTGCTTCTCTCTACTCTCTGAGGGTAGGCAATTTACCTGC 9360
9361 TCCCCAAACCGATGAAACCGGAGCTAACCACTCCAGGCTTTAAGC-----ATTTC 9410
9361 TCCCCAAACCGATGAAACCGGAGCTAACCACTCCAGGCTTTAAGC-----ATTTC 9410
9411 CTGTTT 9470
9421 CTTT 9480
9471 TTTTCTCTCTCTTTTCCCTTCTTTAAATGGTGGCTTCAATCTTAGCCCTAGTACGGCTAGC 9530
9481 TCCTCTTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
9531 TGTGAAGGTCGCTGAGCCGCTGACTGCAGAGAGTGTGATGATGAGCTCTCTGAGAT 9590
9541 TGTGAAGGTCGCTGAGCCGCTGACTGCAGAGAGTGTGATGATGAGCTCTCTGAGAT 9600
9591 CATGT 9595
|||

Db 9601 CAAGT 9605

Search completed: December 11, 2005, 14:05:22
Job time : 30599 secs